

## **Staphylococcus aureus genes and polypeptides**

**[0001]** This application is a divisional of U.S. Application No. 10/138,701, filed May 6, 2002, which is a divisional of U.S. Application No. 09/512,255, filed February 24, 2000 (now U.S. Patent No. 6,403,337, issued June 11, 2002), which is a continuation-in-part of U.S. Application No. 08/781,986, filed January 3, 1997, which claims benefit under 35 U.S.C. § 119(e) of U.S. Provisional Application No. 60/009,861, filed January 5, 1996; U.S. Application No. 09/512,255 is also a continuation-in-part of U.S. Patent Application No. 08/956,171, filed October 20, 1997 (now U.S. Patent No. 6,593,114, issued July 15, 2003), which is a continuation-in-part of U.S. Patent Application No. 08/781,986, filed January 3, 1997, which claims benefit under 35 U.S.C. § 119(e) of U.S. Provisional Application No. 60/009,861, filed January 5, 1996; U.S. Application No. 09/512,255 is also a continuation-in-part of International Application No. PCT/US99/19726, filed August 31, 1999, which claims benefit under 35 U.S.C. § 119(e) of U.S. Provisional Application No. 60/098,964, filed September 1, 1998. Each of the above-listed priority applications is incorporated herein by reference in its entirety.

### **Field of the Invention**

**[0002]** The present invention relates to novel *Staphylococcus aureus* genes (*S. aureus*) nucleic acids and polypeptides. Also provided are vectors, host cells and recombinant methods for producing the same. Further provided are diagnostic methods for detecting *S. aureus* using probes, primers, and antibodies to the *S. aureus* nucleic acids and polypeptides of the present invention. The invention further relates to screening methods for identifying agonists and antagonists of *S. aureus* polypeptide activity and to vaccines using *S. aureus* nucleic acids and polypeptides.

### **Background of the Invention**

**[0003]** The genus *Staphylococcus* includes at least 20 distinct species. (For a review see Novick, R. P., The *Staphylococcus* as a Molecular Genetic System in MOLECULAR BIOLOGY OF THE STAPHYLOCOCCI, 1-37 (R. Novick, Ed., VCH Publishers, New York (1990)). Species differ from one another by 80% or more, by hybridization kinetics, whereas strains within a species are at least 90% identical by the same measure.

[0004] The species *S. aureus*, a gram-positive, facultatively aerobic, clump-forming cocci, is among the most important etiological agents of bacterial infection in humans, as discussed briefly below.

[0005] *Human Health and S. aureus*

[0006] *Staphylococcus aureus* is a ubiquitous pathogen. See, e.g., Mims et al., MEDICAL MICROBIOLOGY (Mosby-Year Book Europe Limited, London, UK 1993). It is an etiological agent of a variety of conditions, ranging in severity from mild to fatal. A few of the more common conditions caused by *S. aureus* infection are burns, cellulitis, eyelid infections, food poisoning, joint infections, neonatal conjunctivitis, osteomyelitis, skin infections, surgical wound infection, scalded skin syndrome and toxic shock syndrome, some of which are described further below.

[0007] *Burns:* Burn wounds generally are sterile initially. However, they generally compromise physical and immune barriers to infection, cause loss of fluid and electrolytes and result in local or general physiological dysfunction. After cooling, contact with viable bacteria results in mixed colonization at the injury site. Infection may be restricted to the non-viable debris on the burn surface ("eschar"), it may progress into full skin infection and invade viable tissue below the eschar and it may reach below the skin, enter the lymphatic and blood circulation and develop into septicemia. *S. aureus* is among the most important pathogens typically found in burn wound infections. It can destroy granulation tissue and produce severe septicemia.

[0008] *Cellulitis:* Cellulitis, an acute infection of the skin that expands from a typically superficial origin to spread below the cutaneous layer, most commonly is caused by *S. aureus* in conjunction with *S. pyogenes*. Cellulitis can lead to systemic infection. In fact, cellulitis can be one aspect of synergistic bacterial gangrene. This condition typically is caused by a mixture of *S. aureus* and microaerophilic *Streptococci*. It causes necrosis and treatment is limited to excision of the necrotic tissue. The condition often is fatal.

[0009] *Eyelid infections:* *S. aureus* is the cause of styes and of "sticky eye" in neonates, among other eye infections. Typically such infections are limited to the surface of the eye, and may occasionally penetrate the surface with more severe consequences.

[0010] *Food poisoning:* Some strains of *S. aureus* produce one or more of five serologically distinct, heat and acid stable enterotoxins that are not destroyed by digestive process of the stomach and small intestine (enterotoxins A-E). Ingestion of the toxin, in sufficient quantities, typically results in severe vomiting, but not diarrhea. The effect does not

require viable bacteria. Although the toxins are known, their mechanism of action is not understood.

[0011] *Joint infections:* *S. aureus* infects bone joints causing diseases such osteomyelitis. See, e.g., R. Cunningham et al., (1996) J. Med. Microbiol. 44:157-164.

[0012] *Osteomyelitis:* *S. aureus* is the most common causative agent of haematogenous osteomyelitis. The disease tends to occur in children and adolescents more than adults and it is associated with non-penetrating injuries to bones. Infection typically occurs in the long end of growing bone, hence its occurrence in physically immature populations. Most often, infection is localized in the vicinity of sprouting capillary loops adjacent to epiphysis growth plates in the end of long, growing bones.

[0013] *Skin infections:* *S. aureus* is the most common pathogen of such minor skin infections as abscesses and boils. Such infections often are resolved by normal host response mechanisms, but they also can develop into severe internal infections. Recurrent infections of the nasal passages plague nasal carriers of *S. aureus*.

[0014] *Surgical Wound Infections:* Surgical wounds often penetrate far into the body. Infection of such wound thus poses a grave risk to the patient. *S. aureus* is the most important causative agent of infections in surgical wounds. *S. aureus* is unusually adept at invading surgical wounds; sutured wounds can be infected by far fewer *S. aureus* cells than are necessary to cause infection in normal skin. Invasion of surgical wound can lead to severe *S. aureus* septicemia. Invasion of the blood stream by *S. aureus* can lead to seeding and infection of internal organs, particularly heart valves and bone, causing systemic diseases, such as endocarditis and osteomyelitis.

[0015] *Scalded Skin Syndrome:* *S. aureus* is responsible for "scalded skin syndrome" (also called toxic epidermal necrosis, Ritter's disease and Lyell's disease). This disease occurs in older children, typically in outbreaks caused by flowering of *S. aureus* strains produce exfoliation(also called scalded skin syndrome toxin). Although the bacteria initially may infect only a minor lesion, the toxin destroys intercellular connections, spreads epidermal layers and allows the infection to penetrate the outer layer of the skin, producing the desquamation that typifies the diseases. Shedding of the outer layer of skin generally reveals normal skin below, but fluid lost in the process can produce severe injury in young children if it is not treated properly.

[0016] *Toxic Shock Syndrome:* Toxic shock syndrome is caused by strains of *S. aureus* that produce the so-called toxic shock syndrome toxin. The disease can be caused by *S. aureus* infection at any site, but it is too often erroneously viewed exclusively as a disease solely of

women who use tampons. The disease involves toxemia and septicemia, and can be fatal.

[0017] *Nocosomial Infections:* In the 1984 National Nocosomial Infection Surveillance Study ("NNIS") *S. aureus* was the most prevalent agent of surgical wound infections in many hospital services, including medicine, surgery, obstetrics, pediatrics and newborns.

*Other Infections:* Other types of infections, risk factors, etc. involving *S. aureus* are discussed in: A. Trilla (1995) *J. Chemotherapy* 3:37-43; F. Espersen (1995) *J. Chemotherapy* 3:11-17; D.E. Craven (1995) *J. Chemotherapy* 3:19-28; J.D. Breen et al. (1995) *Infect. Dis. Clin. North Am.* 9(1):11-24 (each incorporated herein in their entireties).

[0018] *Resistance to drugs of S. aureus strains*

[0019] Prior to the introduction of penicillin the prognosis for patients seriously infected with *S. aureus* was unfavorable. Following the introduction of penicillin in the early 1940s even the worst *S. aureus* infections generally could be treated successfully. The emergence of penicillin-resistant strains of *S. aureus* did not take long, however. Most strains of *S. aureus* encountered in hospital infections today do not respond to penicillin; although, fortunately, this is not the case for *S. aureus* encountered in community infections.

[0020] It is well known now that penicillin-resistant strains of *S. aureus* produce a lactamase which converts penicillin to penicilloic acid, and thereby destroys antibiotic activity. Furthermore, the lactamase gene often is propagated episomally, typically on a plasmid, and often is only one of several genes on an episomal element that, together, confer multidrug resistance.

[0021] Methicillins, introduced in the 1960s, largely overcame the problem of penicillin resistance in *S. aureus*. These compounds conserve the portions of penicillin responsible for antibiotic activity and modify or alter other portions that make penicillin a good substrate for inactivating lactamases. However, methicillin resistance has emerged in *S. aureus*, along with resistance to many other antibiotics effective against this organism, including aminoglycosides, tetracycline, chloramphenicol, macrolides and lincosamides. In fact, methicillin-resistant strains of *S. aureus* generally are multiply drug resistant.

[0022] Methicillin-resistant *S. aureus* (MRSA) has become one of the most important nosocomial pathogens worldwide and poses serious infection control problems. Today, many strains are multiresistant against virtually all antibiotics with the exception of vancomycin-type glycopeptide antibiotics.

[0023] Recent reports that transfer of vancomycin resistance genes from enterococci to *S. aureus* has been observed in the laboratory sustain the fear that MRSA might become resistant



against vancomycin, too, a situation generally considered to result in a public health disaster. MRSA owe their resistance against virtually all  $\beta$ -lactam antibiotics to the expression of an extra penicillin binding protein (PBP) 2a, encoded by the *mecA* gene. This additional very low affinity pbp, which is found exclusively in resistant strains, appears to be the only pbp still functioning in cell wall peptidoglycan synthesis at  $\beta$ -lactam concentrations high enough to saturate the normal set of *S. aureus* pbp 1-4. In 1983 it was shown by insertion mutagenesis using transposon Tn551 that several additional genes independent of *mecA* are needed to sustain the high level of methicillin resistance of MRSA. Interruption of these genes did not influence the resistance level by interfering with PBP2a expression, and were therefore called *fem* (factor essential for expression of methicillin resistance) or *aux* (auxiliary genes).

[0024] Six *fem* genes (*femA*- through *F*) have been described and the minimal number of additional *aux* genes has been estimated to be more than 10. Interference with *femA* and *femB* results in a strong reduction of methicillin resistance, back to sensitivity of strains without PBP2a. The *fem* genes are involved in specific steps of cell wall synthesis. Consequently, inactivation of *fem* encoded factors induce  $\beta$ -lactam hypersensitivity in already sensitive strains. Both *femA* and *femB* have been shown to be involved in peptidoglycan pentaglycine interpeptide bridge formation. FemA is responsible for the formation of glycines 2 and 3, and FemB is responsible for formation of glycines 4 and 5. *S. aureus* may be involved in the formation of a monoglycine mucopeptide precursors. FemC-F influence amidation of the iso-D-glutamic acid residue of the peptidoglycan stem peptide, formation of a minor mucopeptide with L-alanine instead of glycine at position 1 of the interpeptide bridge, perform a yet unknown function, or are involved in an early step of peptidoglycan precursors biosynthesis (addition of L-lysine), respectively.

### Summary of the Invention

[0025] The present invention provides isolated *S. aureus* polynucleotides and polypeptides shown in Table 1 and SEQ ID NO:1 through SEQ ID NO:61. One aspect of the invention provides isolated nucleic acid molecules comprising or alternatively consisting of polynucleotides having a nucleotide sequence selected from the group consisting of: (a) a nucleotide sequence shown in Table 1; (b) a nucleotide sequence encoding any of the amino acid sequences of the polypeptides shown in Table 1; and (c) a nucleotide sequence complementary to any of the nucleotide sequences in (a) or (b). The invention further provides for fragments of the nucleic acid molecules of (a), (b) & (c) above.

**[0026]** Further embodiments of the invention include isolated nucleic acid molecules that comprise, or alternatively consist of, a polynucleotide having a nucleotide sequence at least 90% identical, and more preferably at least 95%, 96%, 97%, 98% or 99% identical, to any of the nucleotide sequences in (a), (b) or (c) above, or a polynucleotide which hybridizes under stringent hybridization conditions to a polynucleotide in (a), (b) or (c) above. Additional nucleic acid embodiments of the invention relate to isolated nucleic acid molecules comprising polynucleotides which encode the amino acid sequences of epitope-bearing portions of a *S. aureus* polypeptide having an amino acid sequence in Table 1, and including but not limited to those epitope-bearing portions shown in Table 4.

**[0027]** The present invention also relates to recombinant vectors, which include the isolated nucleic acid molecules of the present invention, and to host cells containing the recombinant vectors, as well as to methods of making such vectors and host cells. The present invention further relates to the use of these vectors in the production of *S. aureus* polypeptides or peptides by recombinant techniques.

**[0028]** The invention further provides isolated *S. aureus* polypeptides having an amino acid sequence selected from the group consisting of an amino acid sequence of any of the polypeptides described in Table 1 or fragments thereof.

**[0029]** The polypeptides of the present invention also include polypeptides having an amino acid sequence with at least 70% similarity, and more preferably at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% similarity to those described in Table 1, as well as polypeptides having an amino acid sequence at least 70% identical, more preferably at least 75% identical, and still more preferably 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% identical to an amino acid sequence of any of the polypeptides described in Table 1 or fragments thereof. Polynucleotides encoding these polypeptides are also encompassed by the invention.

**[0030]** The present invention further provides a vaccine, preferably a multi-component vaccine comprising one or more of the *S. aureus* polynucleotides or polypeptides described in Table 1, or fragments thereof, together with a pharmaceutically acceptable diluent, carrier, or excipient, wherein the *S. aureus* polypeptide(s) are present in an amount effective to elicit an immune response to members of the *Staphylococcus* genus, or at least *S. aureus*, in an animal. The *S. aureus* polypeptides of the present invention may further be combined with one or more immunogens of one or more other staphylococcal or non-staphylococcal organisms to produce a multi-component vaccine intended to elicit an immunological response against members of the *Staphylococcus* genus and, optionally, one or more non-staphylococcal organisms.

**[0031]** The vaccines of the present invention can be administered in a DNA form, *e.g.*, "naked" DNA, wherein the DNA encodes one or more staphylococcal polypeptides and, optionally, one or more polypeptides of a non-staphylococcal organism. The DNA encoding one or more polypeptides may be constructed such that these polypeptides are expressed as fusion proteins.

**[0032]** The vaccines of the present invention may also be administered as a component of a genetically engineered organism or host cell. Thus, a genetically engineered organism or host cell which expresses one or more *S. aureus* polypeptides may be administered to an animal. For example, such a genetically engineered organism or host cell may contain one or more *S. aureus* polypeptides of the present invention intracellularly, on its cell surface, or in its periplasmic space. Further, such a genetically engineered organism or host cell may secrete one or more *S. aureus* polypeptides. The vaccines of the present invention may also be co-administered to an animal with an immune system modulator (*e.g.*, CD86 and GM-CSF).

**[0033]** The invention also provides a method of inducing an immunological response in an animal to one or more members of the *Staphylococcus* genus, preferably one or more isolates of the *S. aureus* species, comprising administering to the animal a vaccine as described above.

**[0034]** The invention further provides a method of inducing a protective immune response in an animal, sufficient to prevent, attenuate, or control an infection by members of the *Staphylococcus* genus, preferably at least *S. aureus* species, comprising administering to the animal a composition comprising one or more of the polynucleotides or polypeptides described in Table 1, or fragments thereof (*e.g.*, including, but not limited to, fragments which comprise the epitopes shown in Table 4). Further, these polypeptides, or fragments thereof, may be conjugated to another immunogen and/or administered in admixture with an adjuvant.

**[0035]** The invention further relates to antibodies elicited in an animal by the administration of one or more *S. aureus* polypeptides of the present invention and to methods for producing such antibodies and fragments thereof. The invention further relates to recombinant antibodies and fragments thereof and to methods for producing such antibodies and fragments thereof.

**[0036]** The invention also provides diagnostic methods for detecting the expression of the polynucleotides and polypeptides of Table 1 by members of the *Staphylococcus* genus in a biological or environmental sample. One such method involves assaying for the expression of a polynucleotide encoding *S. aureus* polypeptides in a sample from an animal. This expression may be assayed either directly (*e.g.*, by assaying polypeptide levels using antibodies elicited in response to amino acid sequences described in Table 1 or fragments

thereof) or indirectly (*e.g.*, by assaying for antibodies having specificity for amino acid sequences described in Table 1 or fragments thereof). The expression of polynucleotides can also be assayed by detecting the nucleic acids of Table 1. An example of such a method involves the use of the polymerase chain reaction (PCR) to amplify and detect *Staphylococcus* nucleic acid sequences in a biological or environmental sample.

**[0037]** The invention also includes a kit for analyzing samples for the presence of members of the *Staphylococcus* genus in a biological or environmental sample. In a general embodiment, the kit includes at least one polynucleotide probe containing a nucleotide sequence that will specifically hybridize with a *S. aureus* nucleic acid molecule of Table 1 and a suitable container. In a specific embodiment, the kit includes two polynucleotide probes defining an internal region of the *S. aureus* nucleic acid molecule of Table 1, where each probe has one strand containing a 31'-mer-end internal to the region. In a further embodiment, the probes may be useful as primers for polymerase chain reaction amplification.

**[0038]** The present invention also relates to nucleic acid probes having all or part of a nucleotide sequence described in Table 1 which are capable of hybridizing under stringent conditions to *Staphylococcus* nucleic acids. The invention further relates to a method of detecting one or more *Staphylococcus* nucleic acids in a biological sample obtained from an animal, said one or more nucleic acids encoding *Staphylococcus* polypeptides, comprising: (a) contacting the sample with one or more of the above-described nucleic acid probes, under conditions such that hybridization occurs, and (b) detecting hybridization of said one or more probes to the *Staphylococcus* nucleic acid present in the biological sample.

**[0039]** By "biological sample" is intended any biological sample obtained from an individual, body fluid, cell line, tissue culture, or other source which contains *S. aureus* polypeptides or polynucleotides of the invention. As indicated, biological samples include body fluids (such as semen, lymph, sera, plasma, urine, synovial fluid and spinal fluid) which contain the *S. aureus* polypeptides or polynucleotides of the invention, and tissue sources found to contain the expressed *S. aureus* polypeptides shown in Table 1. Methods for obtaining tissue biopsies and body fluids from mammals are well known in the art. Where the biological sample is to include mRNA, a tissue biopsy is the preferred source.

**[0040]** The method(s) provided above may preferably be applied in a diagnostic method and/or kits in which *S. aureus* polynucleotides and/or polypeptides of the invention are attached to a solid support. In one exemplary method, the support may be a "gene chip" or a "biological chip" as described in US Patents 5,837,832, 5,874,219, and 5,856,174. Further, such a gene chip with *S. aureus* polynucleotides of Table 1 attached may be used to diagnose

*S. aureus* infection in a mammal, preferably a human. The US Patents referenced above are incorporated herein by reference in their entirety.

#### Detailed Description

[0041] The present invention relates to recombinant antigenic *S. aureus* polypeptides and fragments thereof. The invention also relates to methods for using these polypeptides to produce immunological responses and to confer immunological protection to disease caused by members of the genus *Staphylococcus*. The invention further relates to nucleic acid sequences which encode antigenic *S. aureus* polypeptides and to methods for detecting *Staphylococcus* nucleic acids and polypeptides in biological samples. The invention also relates to *Staphylococcus* specific antibodies and methods for detecting such antibodies produced in a host animal.

#### [0042] Definitions

[0043] The following definitions are provided to clarify the subject matter which the inventors consider to be the present invention.

[0044] As used herein, the phrase "pathogenic agent" means an agent which causes a disease state or affliction in an animal. Included within this definition, for examples, are bacteria, protozoans, fungi, viruses and metazoan parasites which either produce a disease state or render an animal infected with such an organism susceptible to a disease state (*e.g.*, a secondary infection). Further included are species and strains of the genus *Staphylococcus* which produce disease states in animals.

[0045] As used herein, the term "organism" means any living biological system, including viruses, regardless of whether it is a pathogenic agent.

[0046] As used herein, the term "*Staphylococcus*" means any species or strain of bacteria which is members of the genus *Staphylococcus* regardless of whether they are known pathogenic agents.

[0047] As used herein, the phrase "one or more *S. aureus* polypeptides of the present invention" means the amino acid sequence of one or more of the *S. aureus* polypeptides disclosed in Table 1. These polypeptides may be expressed as fusion proteins wherein the *S. aureus* polypeptides of the present invention are linked to additional amino acid sequences which may be of Staphylococcal or non-Staphylococcal origin. This phrase further includes fragments of the *S. aureus* polypeptides of the present invention.

[0048] As used herein, the phrase "full-length amino acid sequence" and "full-length

polypeptide" refer to an amino acid sequence or polypeptide encoded by a full-length open reading frame (ORF). For purposes of the present invention, polynucleotide ORFs in Table 1 are defined by the corresponding polypeptide sequences of Table 1 encoded by said polynucleotide. Therefore, a polynucleotide ORF is defined at the 5' end by the first base coding for the initiation codon of the corresponding polypeptide sequence of Table 1 and is defined at the 3' end by the last base of the last codon of said polypeptide sequence. As is well known in the art, initiation codons for bacterial species may include, but are not limited to, those encoding Methionine, Valine, or Leucine. As discussed below for polynucleotide fragments, the ORFs of the present invention may be claimed by a 5' and 3' position of a polynucleotide sequence of the present invention wherein the first base of said sequence is position 1.

[0049] As used herein, the phrase "truncated amino acid sequence" and "truncated polypeptide" refer to a sub-sequence of a full-length amino acid sequence or polypeptide. Several criteria may also be used to define the truncated amino acid sequence or polypeptide. For example, a truncated polypeptide may be defined as a mature polypeptide (*e.g.*, a polypeptide which lacks a leader sequence). A truncated polypeptide may also be defined as an amino acid sequence which is a portion of a longer sequence that has been selected for ease of expression in a heterologous system but retains regions which render the polypeptide useful for use in vaccines (*e.g.*, antigenic regions which are expected to elicit a protective immune response).

[0050] Additional definitions are provided throughout the specification.

**[0051] *Explanation of Table 1***

[0052] Table 1 lists the full length *S. aureus* polynucleotide and polypeptide sequences of the present invention. Each polynucleotide and polypeptide sequence is preceded by a gene identifier. Each polynucleotide sequence is followed by at least one polypeptide sequence encoded by said polynucleotide. For some of the sequences of Table 1, a known biological activity and the name of the homolog with similar activity is listed after the gene sequence identifier.

**[0053] *Explanation of Table 2***

[0054] Table 2 lists accession numbers for the closest matching sequences between the polypeptides of the present invention and those available through GenBank and GeneSeq databases. These reference numbers are the database entry numbers commonly used by those

of skill in the art, who will be familiar with their denominations. The descriptions of the nomenclature for GenBank are available from the National Center for Biotechnology Information. Column 1 lists the polynucleotide sequence of the present invention. Column 2 lists the accession number of a “match” gene sequence in GenBank or GeneSeq databases. Column 3 lists the description of the “match” gene sequence. Columns 4 and 5 are the high score and smallest sum probability, respectively, calculated by BLAST. Polypeptides of the present invention that do not share significant identity/similarity with any polypeptide sequences of GenBank and GeneSeq are not represented in Table 2. Polypeptides of the present invention that share significant identity/similarity with more than one of the polypeptides of GenBank and GeneSeq may be represented more than once.

**[0055]      *Explanation of Table 3.***

**[0056]**      The *S. aureus* polypeptides of the present invention may include one or more conservative amino acid substitutions from natural mutations or human manipulation as indicated in Table 3. Changes are preferably of a minor nature, such as conservative amino acid substitutions that do not significantly affect the folding or activity of the protein. Residues from the following groups, as indicated in Table 3, may be substituted for one another: Aromatic, Hydrophobic, Polar, Basic, Acidic, and Small,

**[0057]      *Explanation of Table 4***

**[0058]**      Table 4 lists residues comprising antigenic epitopes of antigenic epitope-bearing fragments present in each of the *S. aureus* polypeptides described in Table 1 as predicted by the inventors using the algorithm of Jameson and Wolf, (1988) Comp. Appl. Biosci. 4:181-186. The Jameson-Wolf antigenic analysis was performed using the computer program PROTEAN (Version 3.11 for the Power MacIntosh, DNASTAR, Inc., 1228 South Park Street Madison, WI). *S. aureus* polypeptides shown in Table 1 may possess one or more antigenic epitopes comprising residues described in Table 4. It will be appreciated that depending on the analytical criteria used to predict antigenic determinants, the exact address of the determinant may vary slightly. The residues and locations shown and described in Table 4 correspond to the amino acid sequences for each polypeptide sequence shown in Table 1 and in the Sequence Listing. Polypeptides of the present invention that do not have antigenic epitopes recognized by the Jameson-Wolf algorithm are not represented in Table 2.

**[0059]      *Nucleic Acid Molecules***

[0060] Sequenced *S. aureus* genomic DNA was obtained from the *S. aureus* strain ISP3. *S. aureus* strain ISP3, has been deposited at the American Type Culture Collection, as a convenience to those of skill in the art. The *S. aureus* strain ISP3 was deposited on 7 April 1998 at the ATCC, 10801 University Blvd. Manassas, VA 20110-2209, and given accession number 202108. As discussed elsewhere herein, polynucleotides of the present invention readily may be obtained by routine application of well known and standard procedures for cloning and sequencing DNA. A wide variety of *S. aureus* strains can be used to prepare *S. aureus* genomic DNA for cloning and for obtaining polynucleotides and polypeptides of the present invention. A wide variety of *S. aureus* strains are available to the public from recognized depository institutions, such as the American Type Culture Collection (ATCC). It is recognized that minor variations in the nucleic acid and amino acid sequence may be expected from *S. aureus* strain to strain. The present invention provides for genes, including both polynucleotides and polypeptides, of the present invention from all the *S. aureus* strains.

[0061] Unless otherwise indicated, all nucleotide sequences determined by sequencing a DNA molecule herein were determined using an automated DNA sequencer (such as the Model 373 from Applied Biosystems, Inc., Foster City, CA), and all amino acid sequences of polypeptides encoded by DNA molecules determined herein were predicted by translation of a DNA sequence determined as above. Therefore, as is known in the art for any DNA sequence determined by this automated approach, any nucleotide sequence determined herein may contain some errors. Nucleotide sequences determined by automation are typically at least about 90% identical, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of the sequenced DNA molecule. The actual sequence can be more precisely determined by other approaches including manual DNA sequencing methods well known in the art. By "nucleotide sequence" of a nucleic acid molecule or polynucleotide is intended to mean either a DNA or RNA sequence. Using the information provided herein, such as the nucleotide sequence in Table 1, a nucleic acid molecule of the present invention encoding a *S. aureus* polypeptide may be obtained using standard cloning and screening procedures, such as those for cloning DNAs using genomic DNA as starting material. See, e.g., Sambrook et al. MOLECULAR CLONING: A LABORATORY MANUAL (Cold Spring Harbor, N.Y. 2nd ed. 1989); Ausubel et al., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY (John Wiley and Sons, N.Y. 1989). Illustrative of the invention, the nucleic acid molecule described in Table 1 was discovered in a DNA library derived from a *S. aureus* ISP3 genomic DNA.



**TABLE 1. Nucleotide and Amino Acid Sequences of *S. aureus* Genes.**

>HGS001, SEQ ID NO:1, fabH, 3-oxoacyl-acyl-carrier protein synthase  
 ATTAAGTGTCAATATTCCTACCTCTGACTTGTAGTTTAAAAAGTAATCTATGTTAAATTAATACCTGGTATTAAAAATTT  
 TATTAAAGAGGTGTTCAACTATGAACGTGGGTATTAAAGGTTTGGTGCATATGCGCCAGAAAAGATTATTGACAATGCC  
 TATTTTGTAGCAATTTTGTAGATACATCTGTGAATGGATTTCTAAGATGACTGGAATTAAAGAAAAGACATTGGGCAGATGA  
 TGATCAAGATACTTCAGATTTAGCATATGAAGCAAGTTTAAAGCAATCGCTGACGCTGGTATTTCAGCCCGAAGATATAG  
 ATATGATAATTGTTGCCACAGCAaCTGGaGATATGCCATTTCCAACGTGCGCAAATATGTTGCAAGAACGTTTAGGGACG  
 GGCAAAGTTGCCTCTATGGATCAACTTGCAGCATGTTCTGGATTTATGTATTCAATGATTACAGCTAAACAATATGTTCA  
 ATCTGGAGATTATCATAACATTTTAGTTGTCTGGTGCAGATAAATATCTAAAATAACAGATTTAACTGACCGTTCTACTG  
 CAGTTCTATTTGGAGATGGTGCAGGTGCGGTTATCATCGGTGAAGTTTCAGATGGCAGAGGTATTATAAGTTATGAAATG  
 GGTTCTGATGGCACAGGTGGTAAACATTTATATTTAGATAAAGATACTGGTAAACTGAAAATGAATGGTCGAGAAGTATT  
 TAAATTTGCTGTTAGAATTATGGGTGATGCATCAACACGTGTAGTTGAAAAAGCGAATTTAACATCAGATGATATAGATT  
 TATTTATTCCTCATCAAGCTAATATTAGAATTATGGAATCAGCTAGAGAACGCTTAGGTATTTCAAAGACAAAATGAGT  
 GTTCTGTGAAATAAATATGGAAATACTTCAGCTGCGTCAATACCTTTAAGTATCGATCAAGAATTAATAAATGGTAAAT  
 CAAAGATGATGATACAATTGTTCTTGTCTGGATTCTGGTGGCGGCTAACTTGGGGCGCAATGACAATAAATGGGGAAAAAT  
 AGGAGGATAACGAATGAGTCAAAATAAAGAGTAGTTATTACAGGTATGGGA

>HGS001, SEQ ID NO:2, FabH, 3-oxoacyl-acyl-carrier protein synthase  
 MNVGIKGFGAYAPEKIIDNAYFEQFLDTSDEWISKMTGIKERHWADDDQDTSDLAYEASLKAIADAGIQPEDIDMIIVAT  
 ATGDMFPFPTVANMLQERLGTGKVASMDQLAACSGFMYSMITAKQVQSGDYHNILVVGADKLSKITDLTDRSTAVLFGDG  
 AGAVIIGEVS DGRGIIISYEMGSDGTGGKHLVLDKDTGKLMNGREVFKFAVRIMGDASTRVVEKANLTSDDIDLFIHQ  
 NIRIMESARERLGISKDKMSVSVNKGNTSAASIPLSIDQELKNGKIKDDDTIVLVGFGGGLTWGAMTIKWK

>HGS002, SEQ ID NO:3, murB, UDP-N-acetylenolpyruvoylglucosamine reductase  
 ATACTAATTCTAATACTTTCTTTTCAATTTTCGCAATGAATTTTAAAATTTGGTATAATACTATATGATATTAAAGACAT  
 GAGAAAGGATGTACTGAGAAGTGATAAATAAAGACATCTATCAAGCTTTACAACAACCTATCCCAAATGAAAAATTTAAA  
 GTTGATGAACCTTTAAAACGATACACTTATACTAAAACAGGTGGTAATGCCGACTTTTACATTACCCCTACTAAAAATGA  
 AGAAGTACAAGCAGTTGTAAATATGCCATCAAAATGAGATTCTGTACATATTTAGGAAATGGCTCAATATTATTA  
 TCCGTGAAGGTGGTATTTCGCGGTATTGTAATTAGTTTATTATCACTAGATCATATCGAAGTATCTGATGATGCGGATAATA  
 GCCGGTAGCGGCGCTGCAATTATTGATGTCTCACGTGTTGCTCGTGATTACGCACCTACTGGCCTTGAATTTGCATGTGG  
 TATTCAGGTTCAATTGGTGGTGCAGTGATATGAATGCTGGCGCTTATGGTGGCGAAGTTAAAGATTGTATAGACTATG  
 CGCTTTGCGTAAACGAACAAGGCTCGTTAATTAAACTTACAACAAAAGAATTAGAGTTAGATTATCGTAATAGCATTATT  
 CAAAAGAACACTTAGTTGTATTAGAAGCTGCATTTACTTTAGCTCCTGGTAAAATGACTGAAATACAAGCTAAAATGGA  
 TGATTTAACAGAACGTAGAGAATCTAAACAACCTTTAGAGTATCCTTCATGTGGTAGTGATTCCAAAGACCGCCTGGTC  
 ATTTTTCAGGTAAATTGATACAAGATTCTAATTTGCAAGGTCACCGTATTGGCGCGCTTGAAGTTTCAACCAACACGCT  
 GGTTTTATGGTAAATGTAGACAATGGAACCTGCTACAGATTATGAAAACCTTATTCATTATGTACAAAAGACCGCTCAAAGA  
 AAAATTTGGCATTGAATTAAATCGTGAAGTTTCGCATTATTGGTGAACATCCAAAGGAATCGTAAGTTAAGGAGCTTTGTC  
 TATGCCTAAAGTTTATGGTTCATTAATCGATACT

>HGS002, SEQ ID NO:4, MurB, UDP-N-acetylenolpyruvoylglucosamine reductase  
 VINKDIYQALQQLIPNEKIKVDEPLKRYTYTKTGGNADFYITPTKNEEVQAVVKYAYQNEIPVTYLNGNSNIIIREGGIR  
 GIVISLLSLDHIEVSDDAI IAGSGAAI IDVSRVARDYALTLGLEFACGIPGSIGGAVYMNAGAYGGEVKDCIDYALCVNEQ  
 GSLIKLTKELELDYRNSIIQKEHLVLEAAFTLAPGKMTEIQAKMDDLTERRESKQPLEYPSGVSFQRPFGHFAGKLI  
 QDSNLQGHRIGGVEVSTKHAGFMVNVNNGTATDYENLIHYVQKTVKEKFGIELNREVRIIGEHPKES

>HGS003, SEQ ID NO:5, fabI, enoyl-acyl-carrier protein reductase  
 AATAGTGTAAATGTATTGACGAATAAAAAAGTTAGTTAAACTGGGATTAGATATTCTATCCGTTAAATTAATTATTAT  
 AAGGAGTTATCTTACATGTTAAATCTTGAACAAAACATATGTCATCATGGGAATCGCTAATAAGCGTAGTATTGCTTT  
 TGGTGTGCGCTAAAGTTTGTAGATCAATTAGGTGCTAAATTAGTATTTACTTACCGTAAAGAACGTAGCCGTAAAGAGCTTG  
 AAAATTTATTAGAACAATTAAATCAACCAGAAGCGCACTTATATCAAATTGATGTTCAAAGCGATGAAGAGGTATTAAAT  
 GGTTTTGTAGCAAATTTGTAAGATGTTGGCAATATTGAGGTGTATATCATTCATTCGATTTGCTAATATGGAAGACTT  
 ACGCGGACGCTTTTCTGAAACTTTCAGTGAAGGCTCTTGTAGCTCAAGACATTAGTTCTTACTCATTAAACAATTGTGG  
 CTCATGAAGCTTAAATAATGCGCAGAAGGTGGTAGCATTTGTGCAACAACATATTTAGGTGGCGAATTTCGCAAGTTCAA  
 AACTATAATGTGATGGGTGTTGCTAAAGCGAGCTTAGAAGCAAATGTTAAATATTTAGCATTAGACTTAGGTCCAGATAA  
 TATTCGCGTTAATGCAATTTTCAGCTAGTCCAATCCGTACATTAAGTGCAAAAGGTGTGGGTGGTTTCAATACAATCTTA  
 AAGAAATCGAAGAGCGTGACCTTTAAAACGTAATGTTGATCAAGTAGAAGTAGGTAAAACCTGCGGCTTACTTTATTAAGT  
 GATTTATCAAGTGGCGTTACAGGTGAAAATATTCATGTAGATAGCGGATTCCACGCAATTAATAATATCATTCAACAGC

TTTGTTACGTTATTATATATGTGAGCAAAGCTTTT

>HGS003, SEQ ID NO:6, FabI, enoyl- acyl-carrier protein reductase

MLNLENKTYVIMGIANKRSIAFGVAKVLDQLGAKLVFTYRKERSRKELEKLLEQLNQPEAHLYQIDVQSDEEVINGFEQI  
GKDVGNIDGVYHSIAFANMEDLRGRFSETSGREGFLLAQDISSYSLTIVAHEAKKLMPEGGSIVATTYLGGEFAVQNYNVM  
GVAKASLEANVKYLALDLGPDNIRVNAISASPIRTLAKVGGFNTILKEIEERAPLKRNVQVEVGKTAAYLLSDLSSG  
VTGENIHVDSGFHAIK

>HGS004, SEQ ID NO:7, murA, UDP-N-acetylglucosamine 1-carboxyvinyltransferase

TAAAAATAATTTTAAAAATAGGGAAATGTAAAGTAATAGGAGTTCTAAGTGGAGGATTTACGATGGATAAAAATAGTAATCAA  
AGGTGGAAATAAATTAACGGGTGAAGTTAAAGTAGAAGGTGCTAAAAATGCAGTATTACCAATATTGACAGCATCTTTAT  
TAGCTTCTGATAAACCGAGCAAATTAGTTAATGTTCCAGCTTTAAGTGATGTAGAAACAATAAATAATGTATTAACAACT  
TTAAATGCTGACGTTACATACAAAAAGGACGAAAATGCTGTTGTCGTTGATGCAACAAAGACTCTAAATGAAGAGGCACC  
ATATGAATATGTTAGTAAATGCGTGCAAGTATTTTAGTTATGGGACCTCTTTTAGCAAGACTAGGACATGCTATTGTTG  
CATTCGCTGGTGGTTGTGCAATTGGAAGTAGACCGATTGAGCAACACATTAAAGGTTTGAAGCTTTAGGCGCAGAAATT  
CATCTTGAAAATGGTAATATTTATGCTAATGCTAAAGATGGATTAAAGGTACATCAATTCATTTAGATTTTCCAAGTGT  
AGGAGCAACACAAAATATTATTATGGCAGCATCATTAGCTAAGGGTAAGACTTTAATTGAAAATGCAGCTAAAGAACCTG  
AAATTGTCGATTTAGCAAACTACATTAATGAAATGGGTGGTAGAATTACTGGTGCCTGACAGACACAATTACAATCAAT  
GGTGTAGAATCATTACATGGTGTAGAATATGCTATCATTCCAGATAGAATTGAAGCAGGCACATTACTAATCGCTGGTGC  
TATAACGCGTGGTGATATTTTGTACGTGGTGCAATCAAAGACATATGGCGAGTTTAGTCTATAAACTAGAAGAAATGG  
CCGTTGAATTGGACTATCAAGAAGATGGTATTCGTGTACGTGCTGAAGGGGAATTACAACCTGTAGACATCAAACTCTA  
CCACATCCTGGATTCCCGACTGATATGCAATCACAATGATGGCATTGTTATTAACGGCAAATGGTCATAAAGCTCGTAAC  
CGAAACTGTTTTTGAACCGTTTTTATGCAATGTTGTCAGAGTTCAAACGTATGAATGCTAATATCAATGTAGAAGTCGTA  
GTGCTAAACTTGAAGGTAAAAAGTCAATTGCAAGGTGCACAAGTTAAAGCGACTGATTAAAGAGCAGCAGCCGCTTAATT  
TTAGCTGGATTAGTTGCTGATGGTAAACAAGCGTTACTGAATTAACGCACCTAGATAGAGGCTATGTTGACTTACACGG  
TAAATTGAAGCAATTAGGTGCAGACATTGAACGTATTAACGATTAATTCAGTAAATTAATATAATGGAGGATTTCAACCA  
TGGAAACAATTTTTGA

>HGS004, SEQ ID NO:8, MurA, UDP-N-acetylglucosamine 1-carboxyvinyltransferase

MDKIVIKGGNKLTGEVKEGAKNAVLPILTASLLASDKPSKLVNVPALSDVETINNVLTTLNADVTKYKDENAVVDATK  
TLNEEAPYEVVSKMRASILVMGP LLARLGHAI VALPGGCAIGSRPIEQHIKGFELGAEIHLNENIYANAKDGLKGTSI  
HLDFFPSVGATQNIIMAA SLAKGKTLIENAAKEPEIVDLANYINEMGGRI TGAGTDTITINGVESLHGVEHAIIPDRIEAG  
TLIIAGAITRGDIFVRGAIKEHMASLVYKLEEMGVELDYQEDGIRVRAEGELQPV DIKTLPHPGFPTDMQSQMMALLLTA  
NGHKVVTETVFENRFMHVAEFKRMNANINVEGRSAKLEKSQLQGAQVKATDLRAAAALILAGLVADGKTSVTELTHLDR  
GYVDLHGKLGADIERIND

>HGS005, SEQ ID NO:9, rho, transcriptional terminator Rho

TTCATGTATTTAAAGGTTGGGGATTAGCATAATGGGATTGTGCTAGCACAGTTATTTATGCATTGTCATGCCTATCTAT  
TACTTACTAACTAAAAATAATGAAATGGGTGTAAACTATATGCCTGAAAGAGAACGTACATCTCCTCAGTATGAATCAT  
TCCACGAATTGTACAAGAACTATACTACCAAGGAACCTCACTCAAAAAGCTAAACTCTTAAGTTGACGAACCATAGTAAA  
TTAAATAAAAAAGAAGCTTGTCTAGCTATTATGGAAGCACAAATGGAAGAAAGATGGTAAGTATTATATGGAAGGTATCTT  
AGATGATATACAACCAGGTGGTTATGGTTTTTAAAGAACAGTGAAGTATTCTAAAGGGGAAAAAGATATTTATATATCTG  
CTAGCCAAATTCGTCGTTTTGAAATTAAACGTGGGGATAAAGTAACTGGGAAAGTTAGAAAACCTAAAGATAACGAAAA  
TATTATGGCTTATTACAAGTTGACTTTGTCAATGACCATAACGCAGAAGAAGTGAAGAAACGTCCGCATTTCCAAGCTTT  
GACACCACCTTTATCCAGATGAGCGTATTAAATTAGAGACAGAAATACAAAATTATTCAACGCGCATCATGGATTTAGTAA  
CACCGATTGGTTTTAGGTCAACGTGGTTTTAATAGTGGCGCCACCTAAAGCAGGTAAACATCGTTATTTAAAGAAATAGCG  
AATGCAATCAGTACGAACAAACAGATGCAAAGCTATTTATTTGTTAGTTGGCGAGCGTCCTGAAGAGGTAACAGATTT  
AGAACGCTCAGTAGAAGCTGCTGAAGTCGTTCAATTCAACGTTTGACGAACCACCAGAACACCATGTTAAAGTAGCTGAAT  
TATTACTTGAACGTGCAAAGCGTTTAGTAGAAAATGGGGAAGATGTCATTATTTAATGGATTCTATAACGAGATTAGCA  
CGCGCTTATAACTTAGTTATTCCACCAAGTGGTCGTACATTATCAGGTGGTTTAGATCCTGCATCTTTACACAAACCAA  
AGCATTCTTCGGTGCAGCGAGAAATATTGAAGCGGGTGAAGTTTAAACAATACTTGCAACTGCATTAGTTGATACGGGTT  
CACGTATGGACGATATGATTTACGAAGAAATTTAAAGGAACAGGTAACATGGAGTTACATTTAGATCGTAAATTTGCTGAA  
CGTCGTATCTTCCCTGCAATTGATATTGGCAGAAGTTCAACGCGTAAAGAAGAATGTTGATAAGTAAATCTGAATTAGA  
CACATTATGGCAATTAAAGAAATCTATTCACTGACTCAACTGACTTTACTGAAAGATTTATTTCGCAACTTAAAGGTCTA  
AGAATAATGAAGATTTCTTCAAGCAGCTACAAAAGTCTGCAGAAGAAAGTACTAAAACGGGTCGACCTATAATTTAATAA  
ACATTATATAGGGGCTTCGCTTTTGAATTAATTACCTTTATAATTACACAGTATTGGGTAAAAACTCACAAATAACTCTG  
TTCCAGATGGTTACAGG

>HGS005, SEQ ID NO:10, Rho, transcriptional terminator Rho  
 MPERERTSPQYESFHLYKNYTTKELTQKAKTLKLTNHSKLNKKELVLAIMEAQMEKDGNYMEGILDDIQGGYGFRLRT  
 VNYSKGEKDIYISASQIRRFEEKRGDKVTGKVRKPKDNEKYVGLQVDFVNDHNAEEVKRPHFQALTPLYPDERIKLET  
 EIQNYSTRIMDLVTPIGLQGRGLIVAPPKAGKTSLLKEIANAISTNKPDAKLFILLVGERPEEVDLERSVEAAEVVHST  
 FDEPPEHHVKVAELLERAKRLVEIGEDVIIIMDSITRLARAYNLVIPPSGRTLSGGLDPASLHKPKAFFGAARNIEAGG  
 SLTILATALVDTGSRMDDMIYEEFKGTGNMELHLDRKLSERRIFPAIDIGRSSTRKEELLISKSELDTLWQLRNLFTDST  
 DFTERFIRKLKRSKNNEFFKQLQKSAEESTKTGRPII

>HGS006, SEQ ID NO:11, rnpA, ribonuclease P protein component  
 GATCTTTTTTTTTCGTTTAAATTAAGAATAAATAGAAATTTATGTTATAAGCTCAATAGAAGTTTAAATATAGCTTCAATA  
 AAAACGATAATAAGCGAGTGATGTTATTGGAAAAAGCTTACCGAATTAAGAAGATGCAGATTTTCAGAGAATATATAAA  
 AAAGTCATTCTGTAGCCAACAGACAATTTGTTGTATACACTTGTAAATAAAGAAATAGACCATTTTCGCTTAGGTAT  
 TAGTGTTCCTAAAAAAGTAGGTAATGCAGTGTTAAGAAACAAGATTAAGAAGCAATACGTGAAAATTTCAAAGTACATA

AGTCGCATATATTGGCCAAAGATATTATTGTAATAGCAAGACAGCCAGCTAAAGATATGACGACTTTACAAATACAGAAT  
 AGTCTTGAGCACGTACTTAAATTTGCCAAAGTTTTTAATAAAAAGATTAAGTAAGGATAGGGTAGGGGAAGGAAAACATT  
 AACCCTCAACACATCCCGAAGTCTTACCTCAGACAAACGTAAGACTGACCTTAGGGTTATAATAACTTACTTT

>HGS006, SEQ ID NO:12, RnpA, ribonuclease P protein component  
 MLLEKAYRIKKNADFQRIYKKGHSVANRQFVVYTCNNKEIDHFRLGISVSKKLGNVLRNLIKRAIRENFVKHSHILAK  
 DIIIVARQPAKDMTTLQIQNSLEHVLKIAKVFNNKKIK

>HGS007M, SEQ ID NO:13, dnaB, replicative DNA helicase  
 CAGCAAAACTGGTGAAGGTGGTAAATGTTTGGGTGAGTAAGTACAAAACAAATGCCGAAGCACTAAAAGCACAAACAT  
 GATATTAATTAAGTAAACGTAATGGAATTTACCAATGGAATTCATTCCCTAGGATATACGAATGTACCTGTTAAAT  
 AGATAAAGAAGTTGAAGGTACAATTCGCGTACACACAGTTGAACAATAAAGTTGGATTGAAATAAGAGGTGAACCATTC  
 ATGGATAGAAATGTATGAGCAAAATCAAATGCCGCTAACAATGAAGCTGAACAGTCTGTCTTAGGTTCAATATTATAGA  
 TCCAGAATTGATTACTACTCAGGAAGTTTTCGTTCTCTGAGTCGTTTTATAGGGTGCCCATCAACATATTTTCCGTG  
 CAATGATGCACTTAAATGAAGATAATAAAGAAATGATGTTGTAACATTGATGGATCAATTATCGACGGAAGGTACGTTG  
 AATGAAGCGGGTGGCCCGCAATATCTTGCAGAGTTATCTACAAATGTACCAACGACGCGAAATGTTTCAAGTATTATACTGA  
 TATCGTTTCTAAGCATGCATTAAAACGTAGATTGATTCAAATGCAGATAGTATTGCCAATGATGGATATAATGATGAAC  
 TTGAAGTAGATGCGATTTAAGTGATGCAGAACGTGCAATTTTAGAGCTATCATCTTCTCGTGAAAGCGATGGCTTTAAA  
 GACATTGAGACGCTTAGGACAAGTGTATGAAACAGCTGAAGAGCTTGATCAAAATAGTGGTCAAACACCAGGTATACC  
 TACAGGATATCGAGATTTAGACCAATGACAGCAGGGTTCAACCGAAATGATTTAATTATCCTTGACGCGCTCCATCTG  
 TAGGTAAGACTGCGTTGCACTTAATATTGCACAAAAGTTGCAACGCATGAAGATATGTATACAGTTGGTATTTTCTCG  
 CTAGAGATGGGTGCTGATCAGTTAGCCACAGTATGATTTGTAGTTCTGGAATGTTGACTCAAACCGCTTAAGAACGGG  
 TACTATGACTGAGGAAGATTGGAGTCGTTTACTATAGCGGTAGGTAAATATCACGTACGAAGATTTTATTGATGATA  
 CACCGGGTATTGCAATTAATGATTTACGTTCTAAATGTGCTCGATTAAAGCAAGAACATGGCTTAGACATGATTGTGATT  
 GACTACTTACAGTTGATTCAAGGTAGTGGTTACGTCGCTCCGATAACAGACAACAGGAAGTTTCTGAAATCTCTCGTAC  
 ATTAAGAGCATTAGCCCGTGAATTAATGTCCAGTTATCGCATTAAGTCAGTTATCTCGTGGTGTGAACAACGACAAG  
 ATAAACGTCCAATGATGAGTGATATTCGTGAATCTGGTTCGATTGAGCAAGATGCCGATATCGTTGCATTCTTATACCGT  
 GATGATTACTATAACCGTGGCGCGGATGAAGATGATGACGATGATGGTGGTTTCGAGCCACAAACGAATGATGAAAACGG  
 TGAAATTGAAATTATCATTGCTAAGCAACGTAACGGTCCAACAGGCACAGTTAAGTTACATTTTATGAAACAATATAATA  
 AATTTACCGATATCGATTATGCACATGCAGATATGATGTAAGAAAGTTTTTCCGTACAATAATCATTAAAGATGATAAAAT  
 TGTACGGTTTTTATTTGTTCTGAACGGGTTG

>HGS007M, SEQ ID NO:14, DnaB, replicative DNA helicase  
 MDRMYEQNQMPHNNEAEQSVLGSIIIDPELINTTQEVLLPESFYRGHQAIFRAMMHLNEDNKEIDVVTLMQDQLSTEGTL  
 NEAGGPQYLAELSTNVPTTRNVQYYTDIVSKHALKRLIQTADSANDGYNDELELDAILSDAERRILELSSRESGDFK  
 DIRDLVGLQVYETAELDQNSGQTPGIPTGYRDLQMTAGFNRLDIIILARPVSVGKTAFALNIAQKVATHEDMYTVGIFS  
 LEMGADQLATRMICSSGNVDSNRLRTGTMTEEDWSRFTIAGVGLSRKIFIDTTPGIRINDLRSKCRRLKQEHGLDMIVI  
 DYQLQIQSGSRASDNRQVEVSEISRTLKALARELKCPVIALSQLSRGVEQRQDKRPMSDIRESGSIEQDADIVAFLYR  
 DDYYNRGGDEDDDDGGFEPQTNDENGEIEIIIAKQRNGPTGTVKLHFMKQYNKFTDIDYAHADMM

>HGS008, SEQ ID NO:15, fabD, malonyl CoA-acyl carrier protein transacylase  
 GTGGTTCCGTATTATTAGGATTGGAAGGTACTGTAGTTAAAGCACACGGTAGTTCAAATGCTAAAGCTTTTTATTCTGCA  
 ATTAGACAAGCGAAATCGCAGGAGAACAAAATATTGTACAAACAATGAAAGAGACTGTAGGTGAATCAAATGAGTAAAA  
 CAGCAATTATTTTCCGGGACAAGGTGCCCAAAAGTTGGTATGGCGCAAGATTTGTTTAAACAACATGATCAAGCAACT  
 GAAATTTTAACTTCAGCAGCGAACACATTAGACTTTGATATTTAGAGACAATGTTTACTGTGAAGAAGGTAAATTTGGG

TGAAACTGAAAACACACAACCAGCTTTATTGACGCATAGTTCGGCATTATTAGCAGCGCTAAAAAATTTGAATCCTGATT  
TTACTATGGGGCATAGTTTAGGTGAATATTCAAGTTTAGTTGACGCTGACGTATTATCATTTGAAGATGCAGTTAAAATT  
GTTAGAAAACGTGGTCAATTAATGGCGCAAGCATTTCTACTGGTGTAGGAAGCATGGCTGCAGTATTGGGATTAGATTT  
TGATAAAGTCGATGAAAATTTGTAAGTCATTATCATCTGATGACAAAATAATTGAACCAGCAAACATTAAATGCCAGGTC  
AAATTGTTGTTTCAGGTCACAAAGCTTTAATTGATGAGCTAGTAGAAAAAGGTAAATCATTAGGTGCAAAACGTGTCATG  
CCTTTAGCAGTATCTGGACCATTCCATTATCGCTAATGAAAGTGATTGAAGAAGATTTTCAAGTTACATTAAATCAATT  
TGAATGGCGTGATGCTAAGTTTCTGTAGTTCAAAATGTAAATGCGCAAGGTGAAACTGACAAAGAAGTAATTAATCTA  
ATATGGTCAAGCAATTATATTCACCAGTACAATTCATTAACCAACAGAATGGCTAATAGACCAAGGTGTTGATCATTTT  
ATTGAAATTGGTCTGGAAAAGTTTATCTGGCTTAATTAAAAAATAAATAGAGATGTTAAGTTAACATCAATTCAAAC  
TTTAGAAGATGTGAAAGGATGGAATGAAATGACTAAGAGTGCTTTAGTAACAGGTGCATCAAGAGGAATTGGACGTAGT  
ATTGCGTTACAATTAGCAGAAGAAGGATATAATGTAGCAGTAACTATGC

>HGS008, SEQ ID NO:16, FabD, malonyl CoA-acyl carrier protein transacylase  
MSKTAIIFPGQGAQKVGMAQDLFNNDQATEILTSAANTLDFDILETMFTDEEGKLGETENTQPALLTHSSALLAALKNL  
NPDFTMGHSLGEYSSLVAADVLSFEDAVKIVRKRQLMAQAFPTGVGSMAAVLGLDFKVD EICKSLSSDDKII EPANIN  
CPGQIVVSGHKALIDELVEKGKSLGAKRVMPLAVSGPFHSSLMKVI EEDFSSYINQFEWRDAKFPVVQNVNAQGETDKEV  
IKSNMVKQLYSPVQFINSTEWLIDQGVDFHIEIGPGKVL SGLIKKINRDVKLTSIQTLEDVKGNWEND

>HGS009, SEQ ID NO:17, alfl, fructose-bisphosphate aldolase  
AAATACACATTTAATCTGCAGTATTTCAATGCATTGACGCTATTTTTTGTATATAATTACTTTGAAAAATACGTGCGTAA  
GCACTCAAGGAGGAACTTTCATGCCTTTAGTTTCAATGAAAGAAATGTTAATTGATGCAAAAGAAAATGGTTATGCGGTA  
GGTCAATACAATTAATAACCTGAAATTCATCAAGCAATTTTAGAAGCGTCACAAGAAGAAAATGCACCTGTAATTTT  
AGGTGTTTCTGAAGGTGCTGCTCGTTACATGAGCGGTTTCTACACAATTGTTAAATGGTTGAAGGGTTAATGCATGACT

TAAACATCACTATTCTGTAGCAATCCATTTAGACCATGGTTCAAGCTTTGAAAAATGTAAAGAAGCTATCGATGCTGGT  
TTCACATCAGTAATGATCGATGCTTCACACAGCCCATTGCAAGAAAACGTAGCAACAAC TAAAAAGTTGTTGAATACGC  
TCATGAAAAAAGGTGTTTCTGTAGAAGCTGAATTAGGTACTGTTGGTGGAACAAGAAGATGATGTTGTAGCAGACGGCATCA  
TTTATGCTGATCCTAAAAGATGTCAAGAACTAGTTGAAAAAACTGGTATTGATGCATTAGCGCCAGCATAGGTTTCAGTT  
CATGGTCCATACAAAGGTGAACCAAAATTAGGATTTAAAGAAATGGAAGAAATCGGTTTATCTACAGGTTTACCATTAGT  
ATTACACGGTGGTACTGGTATCCCGACTAAAGATATCCAAAAGCAATTCCATTTGGTACAGCTAAAATTAACGTAAACA  
CTGAAAACCAAATCGCTTCAGCAAAAGCAGTTCTGTGACGTTTAAATAACGACAAAGAAGTTTACGATCCTCGTAAATAC  
TTAGGACCTGCACGTGAAGCCATCAAAGAAACAGTTAAAGGTAAATTAAGAGTTTCGGTACTTCTAACC CGCTAAATA  
ATTAATATTTAGTCTTTAAGTTATTAATAACGTAGGGATATTAATTTTAAAGAAGCAGACAAAATGGTGTGCTTCTT  
TTTTATGTGCTATAAGTAATAAATAAAACAGTTTGATTTT

>HGS009, SEQ ID NO:18, Alf1, fructose-bisphosphate aldolase  
MPLVSMKEMLIDAKENGYAVGQYNINNLEFTQAILEASQEENAPVILGVSEGAARYMSGFYITVKMVEGLMHDNLNITIPV  
AIHLDHGSSFEKCKEIDAGFTSVMIDASHSPFEENVATTKKVVEYAHEKGVSV EAE LGTVGGQEDDVADGIYADPKE  
CQELVEKTGIDALAPALGVSHPYKGEPLKGFKEEMEEIGLSTGLPLVLHGGTG IPTKDIQKAIPFGTAKINVNTENQIAS  
AKAVRDVLNNDKEVYDPRKYLGPAREAIKETVKGKIKEFGTSNRK

>HGS014, SEQ ID NO:19  
GCTATAATAGGCTAGGTTACAATGAGCTTGCTCATACATATTAATATAATTACAAAAACACGTCGGAGGTACGACATGAT  
TAAAAATACAATTAAAAAATTGATAGAACATAGTATATATACGACTTTTAAATTACTATCAAAATTGCCAAACAAGAATC  
TAATTTATTTTGAAGCTTTTCATGGTAAACAATACAGCGACAACCCCAAAGCATTATGAATACTTAACTGAACATAGC  
GATGCCCAATTAATATGGGGTGTGAAAAAAGGATATGAACACATATTCCAACAGCACAATGTACCATATGTTACAAAGTT  
TTCAATGAAATGGTTT TAGCGATGCCAAGAGCGAAAGCGTGGATGATTAACACACGTACACCAGATTGGTTATATAAAT  
CACC CGAACGACGTACTTACAAACATGGCATGGCAGGCCATTAAAAAAGATTGGTTTGGATATTAGTAACGT TAAATG  
CTAGGAACAAATACTCAAAATTAACCAAGATGGCTTTAAAAAAGAAAGCCAACGGTGGGATTATCTAGTGTCACCTAATCC  
ATATTCGACATCGATATTTCAAATGCATTTTCATGTTAGTCGAGATAAGATTTTGGAAACAGGTTATCCAAGAAATGATA  
AATTATCACATAAACGCAATGATACTGAATATATTAATGGTATTAAGACAAGATTAAATATTCCATTAGATAAAAAAGTG  
ATTATGTACGCGCCAACCTTGGCGTGACGATGAAGCGATTGAGAAGGTTTCATATCAATTTAATGTTAACTTTGATATAGA  
AGCTTTGCGTCAAGCGCTGGATGATGATTATGTTATTTTATTACGCATGCATTATTTAGTTGTGACACGTATTGATGAAC  
ATGATGATTTTGTGAAAGACGTTTCAGATTATGAAGACATTTCCGGATTTTATCTTAATCAGCGATGCGTTAGTTACCGAC  
TACTCATCTGTCTGTTTCGACTTCGGTGATTAATAAGCGTCCGCAAAATTTCTATGCATATGACTTAGATAAATATGGCGA  
TGAGCTTAGAGGTTTTTACATGGATTATAAAAAAGAGTTGCCAGGTCCAATTGTTGAAAAATCAAACAGCACTCATTGATG  
CATTA AAAACAAATCGATGAGACTGCAATGAGTATATTGAAGCACGAACGGTATTTTATCAAAAATCTGTTTCATTAGAA  
GATGGACAAGCGTCACAACGAATTTGCCAAACGATTTTAACTGATAACTTAAAAACAATAAAAAATTATAAATTAATTA  
GTTAAGTGATATAAATAAATAACGAAATGTTTGCTTGTATGTTATTATTTGTGTATGAAA

>HGS014, SEQ ID NO:20

MIKNTIKKLIIEHSIYTTFKLLSKLPKNKLIYFESFHGKQYSDNPKALYEYLTEHSDAQLIWGVKKGYEHIFQQHNPVYVT  
KFSMKWFLAMPRAKAWMINTRTPDWLYKSPRTTYLQTHWGTPLKKIGLDISNVKMLGTNTQNYQDGFKKESQRWDYLVSP  
NPYSTSIFQNAFHVSRDKILETGYPRNDKLSHKRNDTEYINGIKTRLNIPLDKKVIMYAPTWRDDEAIREGSYQFNVNFD  
IEALRQALDDDDYVILLRMHYLVVTRIDEHDDFVKDVSDEYEDISDLYLISDALVTDYSSVMFDFGVLRKPQIFYAYDLDDKY  
GDELRGFYMDYKKELPGPIVENQTALIDALKQIDETANEYIEARTVIFYQKFCSLEDGQASQRICQTIFK

>HGS016, SEQ ID NO:21, murA, UDP-N-acetylglucosamine 1-carboxyvinyltransferase

TGATTTGTAATCAAACTAGATATAATTAATAATGACTTAAAATAATTTTAAAATAGGGAAATGTAAAGTAATAGGAGT  
TCTAAGTGGAGGATTTACGATGGATAAAATAGTAATCAAAGGTGGAAATAAATTAACGGGTGAAGTTAAAGTAGAAGGTG  
CTAAAATGCAGTATTACCAATATTGACAGCATCTTTATTAGCTTCTGATAAACCGAGCAAATTAGTTAATGTTCCAGCT  
TTAAGTGATGTAGAAACAATAAATAATGTATTAACTTTAAATGCTGACGTTACATACAAAAGGACGAAAATGCTGT  
TGTCGTGTGATGCAACAAAGACTCTAAATGAAGAGGCACCATATGAATATGTTAGTAAATGCGTGCAAGTATTTTAGTTA  
TGGGACCTCTTTTAGCAAGACTAGGACATGCTATTGTTGCATTGCTGGTGGTGTGCAATTGGAAGTAGACCGATTGAG  
CAACACATTAAAGGTTTGAAGCTTTAGGCGCAGAAATTCATCTGAAAATGGTAATATTTATGCTAATGCTAAAGATGG  
ATTAAGGTTACATCAATTCATTTAGATTTTCCAAGTGTAGGAGCAACACAAAATATTATTATGGCAGCATCATTAGCTA  
AGGGTAAGACTTTAATTGAAAATGCAGCTAAAGAACCTGAAATTGTCGATTTAGCAAACCTACATTAATGAAATGGGTGGT  
AGAATTACTGGTGTGGTACAGACACAATTACAATCAATGGTGTAGAATCATTACATGGTGTAGAATGCTATCATTCC  
AGATAGAATTGAAGCAGGCACATTACTAATCGCTGGTGTATAACCGCTGGTGATATTTTGTACGTGGTGAATCAAAG  
AACATATGGCGAGTTTAGTCTATAAACTAGAAGAAATGGCGCTTGAATTGGACTATCAAGAAGATGGTATTTCGTGTACGT  
GCTGAAGGGGAATTACACCACTGTAGACATCAAACTCTACCACATCTCGGATTCCCGACTGATATGCAATCACAATGAT  
GGCATTGTTATTAAACGGCAAATGGTCATAAAGTCGTAACCGAACTGTTTTTGAACCGCTTTTATGCAATGTTGCAGAGT  
TCAAACGTATGAATGCTAATATCAATGTAGAAGGTCGTAGTGTAACTTGAAGGTAAAAGTCAATTGCAAGGTGCACAA  
GTTAAAGCGACTGATTTAAGAGCAGCAGCCGCTTAATTTTAGCTGGATTAGTTGCTGATGGTAAAACAAGCGTTACTGA  
ATTAACGCACCTAGATAGAGGCTATGTTGACTTACACGGTAAATTGAAGCAATTAGGTGCAGACATTGAACGTATTAACG  
ATTAATTCAGTAAATTAATATAATGGAGGATTTCAACCATGGAAACAATTTTTGATTATAACCAAATTA

>HGS016, SEQ ID NO:22, MurA, UDP-N-acetylglucosamine 1-carboxyvinyltransferase

MDKIVIKGGNKLTGEVKEGAKNAVLPILTASLLASDKPSKLVNVPALSDVETINNVLTTLNADVTKYKEDENAVVVDTAK  
TLNEEAPYEVVSKMRASILVMGP LLARLGHAI VALPGGAIGSRPIEQHIKGFELGAEIHLNGNIYANAKDGLKGT SI  
HLD FPSVGATQNI IMAASLAKGKTLIENAAKEPEIVDLANY INEMGGRI TGAGTDTITINGVESLHGVEHAI IPDRIEAG  
TL LIAGAITRGDIFVRGAIKEHMASLVYKLEEMGV ELDYQEDGIRVRAEGELQPV DIKTLPHPGFPPTDMQSQMMALLLTA  
NGHKVV TETVFENRFMHVAEFKRMNANINVEGRSAKLEKSQLQGAQVKATDLRAAAALILAGLVADGKTSVTELTHLDR  
GYVDLHGK LKQLGADIERIND

>HGS018, SEQ ID NO:23, dnaJ, DNA ligase

AGAAAAATGGCTCAATCGAACTAGATATTATCTTTAAATCACAAGGGCCAAAACGTTTGTAGCGCAATTTGCACCAATT  
GAAAAAAGGAGGATTAAAGGATGGCTGATTTATCGTCTCGTGTGAACGAGTTACATGATTTATTAAATCAATACAGTTAT  
GAATACTATGTAGAGGATAATCCATCTGTACCAGATAGTGAATATGACAAATTACTTCATGAAGTGAATTTAAATAGAAGA  
GGAGCATCTGAGTATAAGACTGTAGATTCTCCAACAGTTAGAGTTGGCGGTGAAGCCCAAGCCTCTTTCAATAAAGTCA  
ACCATGACACGCCAATGTTAAGTTTAGGGAATGCATTTAATGAGGATGATTTGAGAAAATTCGACCAACGCATACGTGAA  
CAAATGGCAACGTTGAATATATGTGCGAATTTAAATTTGATGGCTTAGCAGTATCATTGAAATATGTTGATGGATACTT  
CGTTCAAGGTTTAAACGTTGGTGTGGAACAACAGGTGAAGATATTACCGAAAATTTAAAAACAATTCATGCGATACCTT  
TGAAAATGAAAGAACCATTAAATGTAGAAGTTCTGTTGGAAGCATATATGCCGAGACGTTTCAATTTTACGATTAAATGAA  
GAAAAAGAAAAAATGATGAGCAGTTATTGCAAAATCCAAGAACGCTGCTGCGGGATCATTAAAGACAGTTAGATTCTAA  
ATTAACGGCAAAACGAAAGCTAAGCGTATTTATATATAGTGTCAATGATTTCACTGATTTCAATGCGCGTTCCGAAAGTG  
AAGCATTAGATGAGTTAGATAAATTAGGTTTACAACGAATAAAAAATAGAGCGCGTGTAATAATATCGATGGTGTGTTTA  
GAGTATATTGAAAAATGGACAAGCCAAAGAGAGTCATTACCTTATGATATTGATGGGATTGTTATTAAAGTTAATGATTT  
AGATCAACAGGATGAGATGGGATTCACACAAAATCTCCTAGATGGGCCATTGCTTATAAAATTTCCAGCTGAGGAAGTAG  
TAACTAAATTTATGATATTGAATTAAGTATTGGACGAACAGGTGTAGTCACACCTACTGCTATTTTGAACAGTAAAA  
GTRGCTGGTACAACGTATCAAGAGCATCTTTGCACAATGAGGATTTAATTCATGACAGAGATATTCGAATTGGTGATAG  
TGTTGTAGTGA AAAAGCAGGTGACATCATACCTGAAGTTGTACGTAGTATCCAGAACGTAGACCTGAGGATGCTGTCA  
CATATCATATGCCAACCCATTGTCCAAGTTGTGGACATGAATTAGTACGTATTGAAGGCGAAGTAGCACTTCGTTGCATT  
AATCCAAAATGCCAAGCAACTTGTGTAAGGATTGATTCATCTTTGATCAAGACAAGCCATGAATATTGATGGTTTAGG  
CACTAAAATTTATTCACAGCTTTATCAAAGCGAATTAATTAAGATGTTGCTGATATTTCTATTTAACAGAGAAGATT  
TATTACCTTTTAGACAGAATGGGGCAGAAAAAGTTGATAATTTATTAGCTGCCATTCAACAAGCTAAGGACAACCTCTTTA  
GAAAAATTTATTATTGTTGCTAGGTATTAGGCATTTAGGTGTTAAAGCGAGCCAAGTGTAGCAGAAAAATATGAAACGAT

AGATCGATTACTAACGGTAACTGAAGCGGAATTAGTAGAAATTCATGATATAGGTGATAAAGTAGCACAACTCTGTAGTTA  
 CTTATTTAGAAAATGAAGATATTCGTGCTTTAATTCAAAAATTAAAAGATAAACATGTTAATATGATTTATAAAGGTATC  
 AAAACATCAGATATTGAAGGACATCCTGAATTTAGTGGTAAAACGATAGTACTGACTGGTAAGYTACATCAAATGACACG  
 CAATGAAGCATCTAAATGGCTTGCATCACAAAGGTGCTAAAGTTACAAGTAGCGTTACTAAAAATACAGATGTCGTTATTG  
 CTGGTGAAGATGCAGGTTCAAAATTAACAAAAGCACAAAGTTTAGGTATTGAAATTTGGACAGAGCAACAAATTTGTAGAT  
 AAGCAAAATGAATTAATAGTTAGAGGGGTATGTCGATGAAGCGTACATTAGTATTATTGATTACAGCTATCTTTATACT  
 CGCTGCTTGTGGTAACCATAAGGATGACCAGGCTGGAAAAGATA

>HGS018, SEQ ID NO:24, DnaJ, DNA ligase

MADLSSRVNELHDLNQQSYEYYVEDNPSVPDSEYDKLLHELIKIEEHPEYKTVDSPTVRVGGEAQASFNKVNHDTPML  
 SLGNFNEDDLRFQDRIREQIGNVEYMCCLKIDGLAVSLKYVDGYFVQGLTRGDGTGEDITENLKIHAIPMKMEPL  
 NVEVRGEAYMPRRSFLRLNEEKEKNDEQLFANPRNAAAGSLRQLDSKLTAKRKLVSFIYSVNDFTDFNARSQSEALDELD  
 KLGFTTNKNRARNVNDIGVLEYIEKWTSQLRESLPYDIDGIVIKVNDLDQQDEMGTQKSPRWAIYKFPAAEVVTKLLDI  
 ELSIGRTGVVTPTAILEPVKVAGTTVSRASLHNEDLIHDRDIRIGDSVVVKAGDIIPEVVRSIPERRPEDAVTYHMPHTH  
 CPSCGHELVRIEIEVALRCINPKCQAQLVEGLIHVFSRQAMNIDGLGTKIIQQLYQSELIKDVADIFYLTEEDLLPLDRM  
 GQKKVDNLLAAIQQAKDNSLENLLFGLGIRHLGVKASQVLAEKYETIDRLLTVTEAEIHDIGDKVAQSVVTYLENED  
 IRALIQLKDKHVNMIYKGIKTSIDIEGHPEFSGKTIIVLTGKLHQMTRNEASKWLASQGAKVTSVTKNTDVIAGEDAGS  
 KLTKAQSLGIEIWTQQFVDKQNELNS

>HGS019, SEQ ID NO:25, mapM, methionine aminopeptidase

TGTCTCACTCACTTTCCAAAATACTAAAGTAACATCTTTAGTATATCAAAGAATTTTGTCTATAATAAGTTATAATTATA  
 TAAAAAAGGAACGGGATAAAAATGATTGTAAAAACAGAAGAATTACAAGCGTTAAAAGAAATTGGATACATATGCGCT  
 AAAGTGCGAATACAAATGCAAGCTGCAACCAAACAGGTATCACTACGAAAGAGCTTGATAATATTGCGAAAGAGTTATT  
 TGAAGAATACGGTGTCTATTTCTGCGCCAATTCATGATGAAAATTTCTCTGGTCAAACGTGTATTAGTGTCAATGAAGAGG  
 TGGCAGATGGGATTCCAAGTAAGCGTGTCTATTCGTGAAGGAGATTTAGTAAATATTGATGTATCGGCTTTGAAGAATGGC  
 TATTATGCAGATACAGGCATTTCTATTTGTCGTTGGAGAATCAGATGATCCAATGAAACAAAAAGTATGTGACGTAGCAAC  
 GATGGCATTGAGAATGCAATTGCAAAAGTAAAACCGGTACTAAGTTAAGTAACATTGGTAAAGCGGTGCATAATACAG  
 CTAGACAAAATGATTTGAAAGTCATTAAAACTTAACAGGTCATGGTGTGGTTTATCATTACATGAAGCACCAGCACAT  
 GTACTTAATTACTTTGATCCAAAAGACAAAACATTATTAAGTGAAGGTATGGTATTAGCTATTGAACCGTTTATCTCATC  
 AAATGCATCATTTGTTACAGAAGGTAAAAATGAATGGGCTTTTGAAACGAGCGATAAAAGTTTTGTTGCTCAAATTGAGC  
 ATACGGTTATCGTGACTAAGGATGGTCCGATTTTAACGACAAAGATTGAAGAAGAATAGTTCAACATATACTAAGACTAA  
 AGTATGAACATCATTTAGTTCGGGAGCCTATTCATATTGGTTTCGGAAGTGTTTTATAATAATTAAGAACACAATCAAT

>HGS019, SEQ ID NO:26, MapM, methionine aminopeptidase

MIVKTEELQALKEIGYICAKVRNTMQAATKPGITTKELDNIAKELFEEYGAISAPIHDENFPQOTCISVNEEVAHGIPS  
 KRVIREDLVNIDVSALKNGYYADTGISFVVGESDDPMKQKVCVDVATMAFENAIKVKPGTKLSNIGKAVHNTARQNDLK  
 VIKNLTGHGVGLSLHEAPAHVLNYPDKDKTLLTEGMVLAIEPFISSNASFVTEGKNEWAFETSDKSFVAQIEHTVIVTK  
 DGPILTTKIEEE

>HGS022-23-24, SEQ ID NO:27, adt, glutamyl-tRNA amidotransferase subunit a, b, and c  
 (operon comprising three ORFs listed below)

TATACAGTTTATATGAAATTAAAGTAGCACCTCATAAATACTTAGATTTTTAATTGGAAATTTGATACAATTTAGTGATG  
 AATGACTTAAAGGAGGCTTTTATTAATGACAAAAGTAACACGTGAAGAAGTTGAGCATATCGCGAATCTTGCAAGACTTC  
 AAATTTCTCCTGAAGAAACGGAAGAAATGGCCAACACATTAGAAAGCATTTTAGATTTTGCAAAACAAAATGATAGCGCT  
 GATACAGAAGGCGTTGAACCTACATATCACGTTTATAGATTTACAAAACGTTTTACGTGAAGATAAAGCAATTAAGGTAT  
 TCCACAAGAATTAGCTTTGAAAAATGCCAAAAGAAACAGAAGATGGACAATTTAAAGTGCCTACAATCATGAATGAGGAGG  
 ACGCGTAAGATGAGCATTCGCTACGAATCGGTTGAGAATTTATTAACTTTAAATAAAAGACAAAAAATCAAACCATCTGA  
 TGTGTAAAGATATATATGATGCAATTGAAGAGACTGATCCAACAATTAAGTCTTTTCTAGCGCTGGATAAAGAAAATG  
 CAATCAAAAAAGCGCAAGAATTGGATGAATTACAAGCAAAAGATCAAATGGATGGCAAATATTGTTGATTCCAATGGGT  
 ATAAAAGATAACATTATTACAAACGGATTAGAAACAACATGTGCAAGTAAATGTTAGAAGTTTTGTGCCAATTTACGA  
 ATCTACTGTAATGGAAAACTACATAATGAAAATGCCGTTTAAATCGGTAAATTAATATGGATGAGTTTGCAATGGGTG  
 GTTCAACAGAAACATCTTATTTCAAAAAACAGTTAAACCCATTTGACCATAAAGCAGTGCCAGGTGGTTTCATCAGGTGGA  
 TCTGCAGCAGCAGTTGCAGCTGGCTTAGTACCATTTAGCTTAGGTTTACAGACACAGGTGGTTCAATTAGACAACCGGCTGC  
 ATATTGTGGCGTTGTCGGTATGAAACCAACATACGGTCGTGTATCTCGATTTGGATTAGTTGCTTTTGCATCTTCATTAG  
 ACCAAATTTGGTCCATTGACTCGAAATGTAAAAGATAATGCAATCGTATTAGAAGCTATTTCTGGTGCAGATGTTAATGAC  
 TCTACAAGTGCACCAGTTGATGATGTAGACTTTACATCTGAAATTTGGTAAAGATATTAAAGGATTAAAAGTTGCATTACC  
 TAAAGAATACTTAGGTGAAGGTGTAGCTGATGACGTAAAAGAAGCAGTTCAAAACGCTGTAGAACTTTAAAATCTTTAG  
 GTGCTGTCTGTTGAGGAAGTATCATTGCCAAATACTAAATTTGGTATTCCATCATATTACGTGATTGCATCATCAGAAGCT

TCGTCAAACCTTTCTCGTTTACGGAATTCGTTATGGTTATCATTCTAAAGAAGCTCATTTCATTAGAAGAATTATATAA  
AATGTCAAGATCTGAAGGTTTCGGTAAAGAAGTAAACGTCGTATTTTCTTAGGTACATTTGCATTAAGTTCAGGTTACT  
ATGATGCTTACTATAAAAAATCTCAAAAAGTTAGAACATTGATTAAGAAATGACTTTGATAAAGTATTCGAAAAATTATGAT  
GTAGTAGTTGGTCCAACAGCGCCTACAACCTGCGTTTAATTTAGGTGAAGAAATTGATGATCCATTAACAATGTATGCCAA  
TGATTTATTAACAACACAGTAAACCTTAGCTGGATTACCTGGTATTTCTGTTCCCTTGTTGGACAATCAAAATGGCCGACCAA  
TCGGTTTACAGTTCAATTGGTAAACCATTCGATGAAAAACGTTATATCGTGTGCGTTATCAATATGAAACACAATACAAT  
TTACATGACGTTTATGAAAAATTATAAGGAGTGGAATCATGCATTTTGAACAGTTATAGGACTTGAAGTTCACGTTAGA  
GTTAAAAACGGACTCAAAAATGTTTCTCCATCACCAGCGCATTTTGGAGCAGAACCTAACTCAAATACAATGTTATCG  
ACTTAGCATATCCAGGTGTCTTACCAGTTGTTAATAAGCGTGCAGTAGACTGGGCAATGCGTGCTGCAATGGCACTAAAT  
ATGGAAATCGCAACAGAATCTAAGTTTGACCGTAAGAATATTTCTATCCAGATAATCCAAAAGCATATCAAAATTTCTCA  
ATTTGATCAACCAATTGGTGAATGGAATATCGATATCGAAGTCGACGGTGAAACAAAACGAATCGGTATTACTCGTC  
TTCACATGGAAGAAGATGCTGGTAAGTCAACACATAAAGGTGAGTATTCATTAGTTGACTTGAACCGTCAAGGTACACCG  
CTAATTGAAATCGTATCTGAACCAGATATTCGTTACCTAAAGAAGCATATGCATATTTAGAAAAATTGCGTTCAATTAT  
TCAATACACTGGTGTATCAGACGTTAAGATGGAAGAGGGATCTTACGTTGTGATGCTAACATCTCTTTACGTCCATATG  
GTCAGAAAAATTTGTTACTAAAGCCGAATTGAAAAACTTAACTCATTAACTATGTACGTAAAGGTTTGAATATGAA  
GAAAAACGCCAAGAAGAAGAAATTGTTAAATGGTGGAGAAATCGGACAAGAAACACGTCGATTTGATGAATCTACAGGTAA  
AACAAATTTAATGCGTGTAAAGAAGGTTCTGATGATTACCGTTACTTCCCAGAGCCTGACATTGTACCTTTATATATTG  
ATGATGCTTGGAAAGAGCGTGTTTCGTCAGACAATTCCTGAATTACCAGATGAACGTAAAGCTAAGTATGTAAATGAATTA  
GGTTTACCTGCATACGATGCACACGTATTAACATTGACTAAAGAAATGTCAGATTTCTTTGAATCAACAAATTGAACACGG  
TGCAGATGTTAAATTAACATCTAATCGTTAATGGGTGGCGTAAACGAATATTTAAATAAAAAATCAAGTAGAATTATTAG  
ATACTAAATTAACACCAGAAAAATTTAGCAGGTATGATTAACTTATCGAAGACGGAACAATGAGCAGTAAATTTGCGAAG  
AAAGTCTTCCCAGATTAGCAGCTAAAGGTGGTAATGCTAAACAGATTATGGAAGATAATGGCTTAGTTCAAATTTCTGA  
TGAAGCAACACTTCTAAATTTGTAATGAAGCATTAGACAATAACGAACAATCAGTTGAAGATTACAAAAATGGTAAAG  
GCAAAGCTATGGGCTTCTTAGTTGGTCAAATTATGAAAGCGTCTAAAGGTCAAGCTAATCCACAATTAGTAAATCAACTA  
TTAAAAACAAGAAATTAGATAAAAGATAATTAAATCATCAAATATGAAGATTAAAAAATAAACCCCTTGATTGCTGACTT  
AGATGCAATCGAGGTTTATTTATATCTATAGAAGTCAAA

>HGS022, SEQ ID NO:28, Adt, glutamyl-tRNA amidotransferase subunit a  
MSIRYESVENLLTLIKDKKIKPSDVVKDIYDAIEETDPTIKSFLALDKENAIKKAQELDELQAKDQMDGKLFIPMGIKD  
NIITNGLETTCAKMLEGFVPIYESTVMEKLNENAVLIGKLNMDDEFAMGGSTETSYFKKTVNPFDHKAVPGSSSGGSAA  
AVAAGLVPFSLGSDTGGSIQPAAYCGVVGMPKPYGRVSRFLVAFASSLDQIGPLTRNVKDNAIVLEAISGADVNDSTS  
APVDDVDFTSEIGKDIKGLKVALPKEYLGEVADDVKEAVQNAVETLKSGLAVVEEVSLPNTKFGIPSYVVIASSEASSN  
LSRFDGIRYGYHSKEAHSLEELYKMSRSEGFGEVKRRIFLGTFFALSSGYDAYYKKSQKVRTLKINDFDKVFENYDVVV  
GPTAPTAFNLGEEIDDPLTMYANDLLTPVNLAGLPGISVPCGQSNRPIGLQFIGKPFDEKTLRVAYQYETQYNLHD  
VYEKL

>HGS023, SEQ ID NO:29, Adt, glutamyl-tRNA amidotransferase subunit b  
MHFETVIGLEVHVELKTDKMFSPSPAHFGEAPNSNTNVIDLAYPGVLPVVNKRVDWAMRAAMALNMEIATESKFDRKN  
YFYDPNPKAYQISQFDQPIGENGYIDIEVDGETKRIGITRLHMEEDAGKSTHKGEYSLVDLNRQGTPLIEIVSEPDIRSP  
KEYAYYLEKLRSIIQYTGVSVDKMEEGSLRCDANISLRPYGQEKFGTKAELKNLNSFNVRKGLEEYKQEEELLNGGE  
IGQETRRFDESTGKTIILMRVKEGSDDYRYFPEPDIVPLYIDDAWKERVQTIPELPDERKAKYVNELGLPAYDAHVLTLT  
KEMSDFFESTIEHGADVKLTSNWLMMGVNEYLNKNQVELLDTKLTPENLAGMIKLIEDGTMSSKIAKVFPELAAGGNA  
KQIMEDNGLVQISDEATLLKFVNEALDNNEQSVEDYKNGKGKAMGLVQIMKASKGQANPQLVNLQLLKQELDKR

>HGS024, SEQ ID NO:30, Adt, glutamyl-tRNA amidotransferase subunit c  
MTKVTRREEVEHIANLARLQISPEETEEMANTLESILDFAKQNDSDTEGVEPTYHVLDLQNVLRDKAIGKIPQELALKN  
AKETEDGQFKVPTIMNEEDA

>HGS025, SEQ ID NO:31, pth, peptidyl-tRNA hydrolase  
CTTACTAAGCTAAAGAATAATGATAATTGATGGCAATGGCGGAAAAATGGATGTTGTCATTATAATAATAAAATGAAACAAT  
TATGTTGGAGGTAAACACGCATGAAATGTATTGTAGGTCTAGGTAATATAGGTAAACGTTTTGAACCTTACAAGACATAAT  
ATCGGCTTTGAAGTCGTTGATTATATTTTAGAGAAAAATAATTTTTCATTAGATAAAACAAAAGTTTAAAGGTGCATATAC  
AATGAACGAATGAACGGCGATAAAGTGTTATTTATCGAACCAATGACAATGATGAATTTGTCAGGTGAAGCAGTTGCAC  
CGATTATGGATTATTACAATGTTAATCCAGAAGATTTAATGTCTTATATGATGATTTAGATTTAGAACAAGGACAAGTT  
CGCTTAAAGACAAAAAGGAAGTGCGGGCGGTCAACAATGGTATGAAGTCAATTAATAAATGCTTGGTACAGACCAATTTAA  
ACGTATTTCGTATTGGTGTGGGAAGACCAACGAATGGTATGACGGTACCTGATTATGTTTACAACGCTTTTCAAATGATG  
AAATGGTAACGATGAAAAAGTTATCGAACACGCAGCACGCGCAATTGAAAAGTTTGTGTAACACATCAGGTTTGAACCAT  
GTTATGAATGAATTTAATGGTGAAGTGAAATAATGACAATATTGACAACGCTTATAAAAAGAAGATAATCATTTTCAAGAC

CTTAATCAGGTATTTGGACAAGCAAACACACTAGTAAC TGGTCTTTC CCGT

>HGS025, SEQ ID NO:32, Pth, peptidyl-trNA hydrolase

MKCI VGLNIGKRFELTRHNIGFEVDYILEKNNFSLDKQKFKGAYTIERMNGDKVLFIEPMTMMNLSGEAVAPIMDYNN  
VNPEDLIVLYDDLDLEQGVRLRQKGSAGGHNGMKSIIKMLGTDQFKRIRIGVGRPTNGMTVPDYVLQRFSDNDEMVTMEK  
VIEHAARAIEK FVETSRFDHVMNEFNGEVK

>HGS026, SEQ ID NO:33

TGATCCGATTATCTTAGTAGGTGCCAATGAAAGTTATGAGCCACGTTGTCGCGCGCACCATATCGTAGCACCTAGTGATA  
ATAATAAGGAGGAATTATAAGTGTGTTGATCAATTAGATATTTGTAGAAGAAAGATACGAACAGTTAAATGAACTGTTAAGT  
GACCCAGATGTTGTAAATGATTAGATAAATTACGTAAATATTTCTAAAGAGCAAGCTGATTTACAAAAAAGCTGATAGTGT  
TTATCGTAACCTATAAAGCTAAAAAAGAAGATTAGCTGATATGAAGAAATGTTAAGTGAGACTGATGATAAAGAAGAAG  
TAGAAATGTTAAAAGAGGAGAGTAATGGTATTAAAGCTGAACTTCCAAATCTTGAAGAAGAGCTTAAATATTATTGATT  
CCTAAAGATCCTAATGATGACAAAGACGTTATTGTAGAAATAAGAGCAGCAGCAGGTGGTGATGAGGCTGCGATTTTTGC  
TGGTGATTTAATGCGTATGTATTCAAAGTATGCTGAATCACAAAGGATTCAAACCTGAAATAGTAGAAGCGTCTGAAAGTG  
ACCATGGTGGTTACAAAGAAATTAGTTTCTCAGTTTCTGGTAATGGCGCGTATAGTAAATTTGAAATTTGAAATGGTGCG  
CACCGCGTTCAACGTGTGCCTGAAACAGAATCAGGTGGACGTTATTCATACTTCAACAGCTACAGTGGCAGTTTACCAGA  
AGTTGAAGATGTAGAAATTGAAATTAGAAATGAAGATTTAAAAATCGACACGTTACGTTCAAGTGGTGCAGGTGGTCAGC  
ACGTAAACACAACCTGACTCTGCAGTACGTATTACCCATTTACCAACTGGTGTCTATTGCAACATCTTCTGAGAAGTCTCAA  
ATTCAAAACCGTGAAAAAGCAATGAAAGTGTTAAAAGCACGTTTATACGATATGAAAGTTCAAGAAGAACACAAAAGTA  
TGCGTCACAACGTAAATCAGCAGTCGGTACTGGTGATCGTTTCAAGACGTTATTCGAACTTATAATTATCCACAAAGCCGTG  
TAACAGACCATCGTATAGGTCTAACGCTTCAAATTTAGGGCAAATTATGGAAGGCCATTTAGAAGAAATTATAGATGCA  
CTGACTTTATCAGAGCAGACAGATAAAATTGAAAGAACTTAATAATGGTGAATTATAAAGAAAAGTTAGATGAAGCAATTC  
ATTTAACACAACAAAAGGGTTTGAACAAACACGAGCTGAATGGTTAATGTTAGATGTATTTCAATGGACCGGTACG

>HGS026, SEQ ID NO:34

VFDQLDIVEERYEQNLNELSDPDVVNDSKLRKYSKEQADLQKTVDVYRNYKAKKEELADIEEMLSETDDKEEVEMLKEE  
SNGIKAELPNLEEELKILLIPKDPNDDKDVIVEIRAAAGGDEAAIFAGDLRMYSKYAESQGFKEIVEASESDHGGYKE  
ISFSVSGNGAYSKLKFENGHRVQRPETESGGRIHTSTATVAVLPEVEDVEIEIRNEDLKIDTYRSSGAGGQHVNTTDS  
AVRITHLPTGVIATSSSEKSIQNRKAMKVLKARLYDMKVQEEQKYASQRKSAVGTGDRSERIRTYNYPQSRVTDHRIG  
LTLQKLGQIMEGHLEEI IDALTLSEQTDKLKELNNGEL

>HGS028, SEQ ID NO:35

ATTTCTTAACATTGTTATTTAAACAAAATTATGTTAAAATTTAGCATTATAAAAGATGCAAATCAATGACTTGAATTGAAA  
TATAAATAGGAGCGAATGCTATGGAATTATCAGAAATCAAACGAAATATAGATAAGTATAATCAAGATTTAACACAAATT  
AGGGGGTCTCTTGACTTAGAGAACAAAGAACTAATATTCAAGAATATGAAGAAATGATGGCAGAACCTAATTTTGGGA  
TAACCAAAGCAAGCGCAAGATATTATAGATAAAAAATATGCGTTAAAAGCAATAGTTAATGGTTATAAAACACTACAAG  
CAGAAGTAGATGACATGCTACTTGGGATTTATTACAAGAAGAAATTTGATGAAGAAATGAAGAAGACTTAGAGCAA  
GAGGTCATTAATTTTAAGGCTAAAGTGGATGAATACGAATTTGCAATTATTATTAGATGGGCCTCAGCATGCCAAATACGC  
AATTCTAGAGTTACATCCTGGTGCAGGTGGCAGGAGTCTCAAGATTGGGCTAATATGCTATTTAGAAATGTATCAACGTT  
ATTGTGAGAAGAAAGGCTTTAAAGTTGAACTGTTGATTATCTACCTGGGGATGAAGCGGGGATTTAAAGTGTAACATTG  
CTCATCAAAGGGCATAATGCTTATGGTTATTTAAAAGCTGAAAAAGGTGTACACCGACTAGTACGAATTTCTCCATTGTA  
TTCATCAGGACGTCGTCATACATCATTTGCATCATGCGACGTTATTCCAGATTTTAATAATGATGAAATAGAGATTGAAA  
TCAATCCGGATGATATTACAGTTGATACATTAGAGCTTCTGGTGCAGGTGGTCAGCATATTAACAAAACCTGAATCGGCA  
ATACGAATTACCCACCACCCCTCAGGTATAGTTGTTAATAACCAAATGAACGTTCTCAAATTTAAAACCGTGAAGCAGC  
TATGAAAATGTTAAAGTCTAAATTATATCAATTTAAATTTGGAAGAGCAGGCACGTGAAATGGCTGAAATTCGTGGCGAAC  
AAAAAGAAATCGGCTGGGGAAGCCAAATTAGATCATATGTTTCCATCCATACTCAATGGTGAAAGATCATCGTACGAAC  
GAAGAAACAGGTAAGGTTGATGCAGTGATGGATGGAGACATTGGACCATTATCGAATCATATTTAAGACAGACAATGTC  
GCACGATTAATATATATTTTAAAACCGAGGCTCTAAAAGGGCGTGGGTTTTTGGTTTTTTTTAAAGGTAGCTAAATAAATT  
GTAATTAGATTTTGAATATGATTTGTTTATGAA

>HGS028, SEQ ID NO:36

MELSEIKRNIDKYNQDLTQIRGSLDLENKETNIQIEYEMMAEPNFWDNQTKAQDIIDKNNALKAIVNGYKTLQAEVDDMD  
ATVDLLQEEFDEEMKEDLEQEVINFKAKVDEYELQLLLDGPHDANNAILELHPGAGGTESQDWANMLFRMYQRYCEKKGF  
KVETVDYLPGEAGIKSVTLLIKGHNAYGYLKAEGVHRLVRI SPFDSSGRRHTSFASCDVIPDFNNDIEIEINPDDIT  
VDTFRASAGGQHINKTESAIRITHHPGIVVNNQNERSQIKNREAAMKMLKSKLYQLKLEEQAEMAEIRGEQKEIGWG  
SQIRSYVFHPYSVMVDHRTNEETGKVDAVMDGDIGPFIESYLRQTM SHD



>HGS030, SEQ ID NO:37, Tmk, thymidylate kinase  
AATAACTGAAAATATGATAGAATTGGTAAATGAATATCTGGAACTGGAATGATAGTTGAAGGAATTAATAAATAATAAAAA  
TTTTAGTTGAGGATGAATAAAATGTCAGCTTTTATAACTTTTGAGGGCCCAGAAAGGCTCTGGAAAAACAACGTGAATTAA  
TGAAAGTTTACCATAGATTAGTAAAGATTATGATGTCATTAGTACTAGAGAACCAGGTGGTGTTCCTACTGGTGAAGAAA  
TACGTAAAATTGTATTAGAAGGCAATGATATGGACATTAGAACTGAAGCAATGTTATTGCTGCATCTAGAAGAGAACAT  
CTTGATTAAAGGTCATACCAGCTTTAAAGAAGGTAAGGTTGTGTGTGTGATCGCTATATCGATAGTTTCATTAGCTTA  
TCAAGGTTATGCTAGAGGGATTGGCGTTGAAGAAGTAAGAGCATTAAACGAATTTGCAATAAATGGATTATATCCAGACT  
TGACGATTTATTTAAATGTTAGTGCTGAAGTAGGTCGCGAACGTATTATTAATAAATTCAGAGATCAAAATAGATTAGAT  
CAAGAAGATTTAAAGTTTCACGAAAAAGTAATTGAAGGTTACCAAGAAATCATTCATAATGAATCACAACGGTTCAAAAG  
CGTTAATGCAGATCAACCTCTTGAAAATGTTGTGTAAGACACGTATCAAACTATCATCAAAATTTAGAAAAGATATGAT  
ATAATTGTTAGAAGAGGTGTTATAAAATGAAAATGATTATAGCGATCGTACAAGATCAAGATAGTCAGGAACCTGCAGAT  
CAACTTGTTAAATAAATCTTTAGAGCAACAAAATTTGGCAA

>HGS030, SEQ ID NO:38, tmk, thymidylate kinase  
MSAFITFEGPEGSGKTTVINEVYHRLVKDYDVIMTREPGGVPTGEEIRKIVLEGNDMDIRTEAMLFAASRREHLVLKVIP  
ALKEGKVVLCDRYIDSSLAYQGYARGIGVEEVRALNEFAINGLYPDLTIYLNVS AEVGRERI IKNSRDQNRDLQEDLKFH  
EKVIEGYQEIIHNESQRFKSVNADQPLENVVEDTYQTI IKYLEKI

>HGS031, SEQ ID NO:39, PyrH, uridylate kinase  
AATGTTGCTTTTATTAATAATGTAATCATTTCTAATAAAACGACAACTGTGTCTTCTTTACTTGTATATGTTACATATATTC  
ACGATAGAGAGGATAAGAAAATGGCTCAAATTTCTAAATATAAACGTGTAGTTTGTAACTAAGTGGTGAAGCGTTAGCT  
GGAGAAAAAGGATTTGGCATAAAATCCAGTAATTATTAAGAGTTGCTGAGCAAGTGGCTGAAGTTGCTAAAATGGACTG  
TGAAATCGCAGTAATCGTTGGTGGCGGAAACATTTGGAGAGGTA AAAACAGGTAGTGAAGTATGGATGGACCGTGGAACTG  
CTGATTACATGGGTATGCTTGCAACTGTAATGAATGCCTTAGCATTACAAGATAGTTTAGAACAATTGGATTGTGATACA  
CGAGTATTAACATCTATTGAAATGAAGCAAGTGGCTGAACCTTATATTCGTCGTCGTGCAATTAGACACTTAGAAAAAGAA  
ACGCGTAGTTATTTTGTCTGCAGGTATTGGAAACCCATACTTCTCTACAGATACTACAGCGGCATTACGTGCTGCAGAAG  
TTGAAGCAGATGTTATTTAATGGGCAAAAATAATGTAGATGGTGTATATTCTGCAGATCCTAAAGTAAACAAAGATGCG  
GTAAAATATGAACATTTAACGCATATTCAAATGCTTCAAGAAGGTTTACAAGTAATGGATTCAACAGCATCTCTATTCTG  
TATGGATAATAACATTCCGTTAACTGTTTCTCTATTATGGAAGAAGGAAATATTAAACGTGCTGTTATGGGTGAAAAGA  
TAGGTACGTTAATTACAAAATAAATTTAGAGGTGTAAAATAATGAGTGACATTATTAATGAACTAAATCAAGAATGCAA  
AAATCAATCGAAAGCTTATCACGTGAATTAGCTAACATCAGT

>HGS031, SEQ ID NO:40, pyrH, uridylate kinase  
MAQISKYKRVVLKLSGEALAGEKGFGINPVIIKSVAEQVAEVAKMDCEIAVIVGGGNIWRGKTGSDLGMDRGTADYMGML  
ATVMNALALQDSLEQLDCDTRVLTSEIMKQVAEPYIRRAIRHLEKKRVVIFAAGIGNPYFSTDTTAALRAAEVEADVIL  
MGKNNVDGVYSADPKVNKDAVKYEHLLTHIQMLQEGLOQVMDSTASSFCMDNNIPLTVFSIMEEGNIKRAVMGEKIGTLITK

>HGS032, SEQ ID NO:41  
GATAGCATCCATGTATAGTGATAGTATTTACAACAATTATTATAATACTATTTAGTTAAGTAGAGAAATAGTTAAACATT  
TGAAAGTGTGGTTTAAATGGAATGTCAGCAATAGGAACAGTTTAAAGAACATGTAAAGAACTTTTATTAAATTCAAAGA  
CTGGCTCAGTTTCAAGTTAAAATTATCAATCATAGTAACATTTAGGTGTGGCTTGGAATTAATTAACCTGTTATGCA  
AATTATGGTTTACTGGATGGTTTTTGGATTAGGAATAAGAAGTAATGCACCAATTCATGGTGTACCTTTTGTATTATGGT  
TATTGGTTGGTATCAGTATGTGGTTCTTCATCAACCAAGGTATTTAGAAAGTACTAAAGCAATTACACAAAAGTTAAT  
CAAGTATCGAAAATGAACCTCCCGTTATCGATAATACCGACATATATTGTGACAAGTAGATTTTATGGACATTTAGGCTT  
ACTTTTACTTGTGATAATTGCATGTATGTTACTGGTATTTATCCATCAATACATATCATTCAATTATTGATATATGTAC  
CGTTTTGTTTTTCTTAACCTGCCTCGGTGACGTTATTAACATCAACACTCGGTGTGTTAGTTAGAGATACACAAATGTTA  
ATGCAAGCAATATTAAGAATATTATTTTACTTTTCACCAATTTGTGGCTACCAAGAACCATGGTATCAGTGGTTTAAAT  
TCATGAAATGATGAAATATAATCCAGTTTACTTTATTGCTGAATCATACCGTGCAGCAATTTTATATCAGCAATGGTATT  
TCATGGATCATTGGAATTAATGTTATACAATTTCCGTATTGTTGCCATTTCTTTGCAATTGGTGGCTACTTACACATG  
AAATATAGAGATCAATTTGCAGACTTCTTGTAATATATTTATATGACGAAACCCCGCTAACCATTAATAAATGGAAGTGG  
GGTTCATTTTGTGTTATAATTTAAGTAAATAACATATTAAGTTGGTGTATTAT

>HGS032, SEQ ID NO:42  
MSAIGTVFKEHVKNFYLIQRLAQFQVKI INHSNYLGVAVELINPVMQIMVYWMVFLGIRSNAPIHGVPFVYWL LVGISM  
WFFINQIGILEGTAITQKFNQVSKMNFPLSI IPTYIVTSRFGHLGLLLLVI IACMFTGIYPSIHI IQLLIYVPFCFFLT  
ASVTLTSTLGLVLRDQTQMLMQAILRLIFYFSPILWLPKNHGISGLIHEMMKYNPVYFIAESYRAAILYHEWYFMDHWKL  
MLYNFGIVAIFFAIGAYLHMKYRDQFADFL

>HGS033, SEQ ID NO:43

TAACAAAATCTTCTATACACTTTACAACAGGTTTTAAATTTAACAACGTTGAGTAGTATATTATAATCTAGATAAATG  
TGAATAAGGAAGGTCTACAAATGAACGTTTCGGTAAACATTAATAATGTAACAAAAGAATATCGTATTTATCGTACAAAT  
AAAGAACGTATGAAAGATGCGCTCATTCCCAAACATAAAAACAAAACATTTTTTCGCTTAGATGACATTAGTTTAAAGC  
ATATGAAGGTGACGTCATAGGGCTTGTGGCATCAATGGTCCGGCAAATCAACGTTGAGCAATATCATTGGCGGTTCTT  
TGTCGCCCTACTGTTGGCAAAGTGGATCGTAATGGTGAAGTCAGCGTTATCGCAATTAGTGCTGGCTTGAGTGGACAACCTT  
ACAGGGATTGAAAATATCGAATTTAAATGTATGTATGGGCTTTAAGCGAAAAGAAATTAAGCGATGACACCTAAGAT  
TATTGAATTTAGTGAACCTGGTGAGTTTATTATCAACCAGTTAAAAAGTATTCAAGTGGTATGCGTGCAAACTTGGTT  
TTTCAATTAATATCACAGTTAATCCAGATATCTTAGTCATTGACGAAGCTTTATCTGTAGGTGACCAACTTTTGCACAA  
AAATGTTTAGATAAAATTTACGAGTTTAAAGAGCAAAACAAAACCATCTTTTTTCGTTAGTCATAACTTAGGACAAGTGAG  
ACAATTTTGTACTAAGATTGCTTGGATTGAAGGCGGAAAGTTAAAAAGATTACGGTGAACCTTGATGATGTATTACCTAAAT  
ATGAAGCTTTTCTTAAACGATTTTAAAAAGAAATCCAAAGCCGAACAAAAGAATTTAGAAACAACTCGATGAGTCCCGC  
TTCGTTATTAAATAAACCGAAAAACCGAGAACTCCATTAAAGGATTTCTCTCGGTTTATTTTTGTGCATCATGATTATT  
TCGCCTTTTTTATTTTCTTTTTTGCTTTGGCTATT

>HGS033, SEQ ID NO:44

MNVSVNIKVNKEYRIYRTNKERMKDALIPKHKNKTFALDDISLKAYEGDVIGLVGINGSKSTLSNIIGGSLSPVVGK  
VDRNGEVSVIAISAGLSGQLTGIEINIEFKMLCMGFKRKEIKAMTPKIIIEFSELGEFIYQPVKKYSSGMRKLGFSINITV  
NPDILVIDEALSVGDQTFQKCLDKIYEFKEQNKTIFVSHNLGQVRQFCTKIAWIEGKLDYGELEDDVLPKYEAFLND  
FKKSKAEQKEFRNKLEDSRFVIK

>HGS034, SEQ ID NO:45

ATAAGGTGAAGACACATAAAACAATATATCTTAGTAAGCATGCAACACTCTTTTTTGTATTATTCATAACAACAAAAAGA  
ATTAAAGGAGGAGTCTTATTATGGCTCGATTTCAGAGGTTCAAACCTGGAAAAAATCTCGTCGTTTAGGTATCTCTTAAAGC  
GGTACTGGTAAAGAATTAGAAAAACGTCCTTACGCACCAGGACAACATGGTCCAAACCAACGTAAAAAATTATCAGAATA  
TGGTTTACAATTACGTGAAAAACAAAAATTACGTTACTTATATGGAATGACTGAAAGACAATTCGTAACACATTTGACA  
TCGCTGGTAAAAAATTCGGTGATACACGGTGAACCTTCATGATCTTATTAGCAAGTCGTTTAGACGCTGTTGTTTATTCA  
TTAGGTTTAGCTCGTACTCGTCGTCAGCACGTCAATTAGTTAACCACGGTCATATCTTAGTAGATGGTAAACGTGTTGA  
TATTCCATCTTATCTGTAAACCTGGTCAAACAATTTTCACTTCGTGAAAAATCTCAAAAAATTAAACATCATCGTTGAAT  
CAGTTGAAATCAACAATTTTCGTACCTGAGTACTTAACTTTGATGCTGACAGCTTAAGTGGTACTTTTCGTACGTTTACCA  
GAACGTAGCGAATTACCTGCTGAAATTAACGAACAATTAATCCGTTGAGTACTACTCAAGATAATACGGTCAATACCAAC  
ACCCACAATTGTGGGTGT

>HGS034, SEQ ID NO:46

MARFRGSNWKKSRRRLGISLGTGKELEKRPYAPGQHPNQKRLSEYGLQLREKQKLRYLYGMTERQFRNTFDIAGKKFG  
VHGENFMILLASRLDAVVYSLGLARTRRQARQLVNHGILVDGKRVDIPSYSVKPGQTISVREKSQKLNIIVESVEINN  
VPEYLNFDADSLTGTFVRLPERSELPAEINEQLIR

>HGS036, SEQ ID NO:47

TGTTGATTGCACCTGCTTCAGTCATTGCTATAACTATTTTAATTTTAAATTTAACCGGTGATGCCTAAGAGATAGATTG  
CTGAAACAACGGGTGAATATGATGAGTCTCATTGATATACAAATTTAACAATAAAGAATACTAGTGAGAAATCTCTTA  
TTAAAGGGATTGATTGAAAATTTTTTAGTCAACAGATTAATGCCCTTGATTGGAGAGAGCGGCGCTGGAAAAAGTTTGATT  
GCTAAAGCTTTACTTGAATATTTACCATTGATTAAAGCTGCACGTATGATTTCGTACCAATTTGATGGGGAAAATGTTAG  
TAGATTGAGTCAATATATGGTCATACAATTGGCTATATTTCTCAAAATTATGCAGAAAGTTTTAACGACCATACTAAAT  
TAGGTAAACAGTTAACTGCGATTTATCGTAAGCATTATAAAGGTAGTAAAGAAGAGGCTTTGTCCAAAGTTGATAAGGCT  
TTGTGCTGGGTTAATTTACAAAGCAAAGATATATTAATAAATATAGTTTCCAACCTTTCTGGGGGCCAACTTGAACGCGT  
ATACATAGCAAGCGTTCTCATGTTGGAGCCTAAATTAATCATTGCAGACGAACCAAGTTGCATCATTGGATGCTTTGAACG  
GTAATCAAGTGATGGATTTATTACAGCATATTGTATTAGAACATGGTCAAACATTATTTATTATCACACATAACTTAAGT  
CATGTATTGAAATATGTGAGTACATTATGTTTTAAAGAAGGTCAAATCATTGAACGAGGTAATATTAATCATTTCAA  
GTATGAGCATTTGCATCCGTATACTGAACGTCTAATTAATATAGAACACAATTAAAGAGGGATTACTATGATTGAGTTA  
AAACATGTGACTTTTGGTTATAATAAAAAGCAGATGGTGTCTACAAGATATCAATATTACTATACCTGATGGAGAAAATGT  
TGGTATTTTAGGCGAAAGTG

>HGS036, SEQ ID NO:48

MMSLIDIQNLTIKNTSEKSLIKGIDLKIFSQQINALIGESGAGKSLLAKALLEYLFPDLSTYDSYQFDGENVSRLSQYY  
GHTIGYISQNYAESFNDHTKLKQLTAIYRKHYKGSKEEALSKVDKALSWNLQSKDILNKYSFQLSGGQLERVYIASVL  
MLEPKLIIADEPVASLDALNGNQVMDLLQHIVLEHGQTLFIITHNLSHVLKYCQYIYVLKEGQIIERGNINHFKYEHLP  
YTERLIKYRTQLKRDYYD

>HGS040, SEQ ID NO:49

GATGATATTTTAATTACAGAAAATGGTTGTCAAGTCTTTACTAAATGCACAAAAGACCTTATAGTTTTAACATAAGCGTG  
TAAAATGAGGAGGAACTGAATGATTTTCGGTTAATGATTTTAAACAGGTTTAAACATTTCTGTTGATAACGCTATTTGG  
AAAGTTATAGACTTCCAACATGTAAAGCCTGGTAAAGGTTTCAGCATTCGTTTCGTTCAAATTTACGTAATTTAAGAAGTGG

TGCAATTCAAGAGAAAACGTTTAGAGCTGGTGAAAAAGTTGAACCAGCAATGATTGAAAATCGTCGCATGCAATATTTAT  
ATGCTGACGGRGATAATCATGTATTTATGGATAAATGAAAGCTTTGAACAAACAGAACCTTCAAGTGATTACTTAAAGAA  
GAATTGAATTACTTAAAGGAAGGTATGGAAGTACAAATTCAAACATACGAAGGTGAAACTATCGGTGTTGAAATTACCTAA  
AACTGTTGAATTAACAGTAACTGAAACAGAACCTGGTATTAAAGGTGATACTGCAACTGGTGCCACTAAATCGGCAACTG  
TTGAAACTGGTTATACATTAAATGTACCTTTATTTGTAACGAAGGTGACGTTTTAATTATCAACACTGGTGATGGAAGC  
TACATTTCAAGAGGATAATCTCTAATTTGTTAAACAAATAGCTTGTATTCACTATACTGATTTAACGTAAGANATCTAAA  
TAAGTCTCATAAAGCTATTGCCTAAAATGATTATAGGTTA

>HGS040, SEQ ID NO:50

MISVNDPKTGLTISVDNAIWKVIDFQHVKPGKGSFVRSKLRLNLTGAIQEKTFRAGEKVEPAMIENRRMQYLYADGDNH  
VFMNDNESFEQTELSSDYLKEELNYLKEGMEVQIQTYEGETIGVELPKTVELTVTETEPGIKGDATGATKSAATVETGYTL  
NVPLFVNEGDLVIINTGDGSIISR

>168153/168339, SEQ ID NO:51, (operon comprising ORFs for five polypeptides listed below)

TTAGGATGTAAGAAAGTTCCAGTGCAAGAAATCCATGAAACACAATATTCAATTAGTACATGGCAACATAAAGTTCCATTTGGT  
GTGTGGTGGGAAACGTTACAACAAGAATCATCGCTTGCCATGGACTACTGAGACAAGACAAGAAGCGCCATTTATTACAATGTGT  
CATGGTGATACAGAACATATTTGTATACAAAAGATTAGGCCAAGCACATTTTCAAGTATGGGAAAAGGTTGTGCGCAAGTTAT  
AGTGGTTGTTGTTCTGTAGAGAGAATGGCACAAAGTACATATCTGTCTTCTCAACAAGATGTACTCATGAAGTATCGCCA  
TTGAGTTATAAGGAAATGAAGCGGTTGTTTCATAAAGGGGAAACTGTGCCAGCAGGTGTGACACGCTTTAATATTTTCAGGACGA  
TGTCTTAATCTTCAAGTACCCTGGCATTACTTAAACAAGATGATGATGTTGAACAATGCGCAATTTGGAAGCAGTTTTCAGCAG  
ATAAGTTTGGCAAATGATGATGCTATACGTGAAAAGTATACCTTGGTGAGCAATAGTTTACTGTGATGTTGAGGGAATATGA  
TGATTTAGCGTATTGATAGCGAAAATATAATAAAACAATATAGTGTGGAGAATTTTGTATTTTATAAATATTGAAGTTCTCC  
ATTTTGTATTTTGCATATAAAAATTAATAAAAATAAGGTATTAAGGTAAAGTATAAATTTTAAATAAATGGGGAATGAGTA  
TGAGTCAATTATAGGAAAATAGCAATTTGGATAGGCATCTGAGTCAATATATTTTAGTGTCTGTTTTTGTAGGATGATAT  
CTATTAATATTGCTGGAGGATCTGATTACGAAACAATTTTTTTATAGGATTAATATTGGCTCTTTTTCACTGTTTTACCAACCA  
TCCTTACTGCGATTTATATGGAAGTTACTCTGTAATCGGAGGTGCACTTTTTATTGTTTATGCTATTATTGCACTGTGTTTTAT  
ATAATTTCTTTCTCAATTTTATGGCTGATTGGTGGTATTTTGTCTGATTGGAATAAATACTCAAAAGATGAATCGACAGACG  
AAAAAGAAAAGTTGATATTGAAAAGTACAGAGAATCAATTTGAATCTAAAGATAAATACTCAAAAGAATAAGAGAATATTTAA  
GGTAAAGTATAAATTTTAAATAAATGGGGAATAGACATGGAAAAAATGTAGAAAAATCATTCAATAAGATAGGTTTATATTTT  
CAAATAGCTTATATAGTACTCATGGCTATAACTTTATGTGGGTTTGTAAATTTGCTATGGCAATATTTTCGGCCTTTTCTATTTA  
TTATCAGGTAGCAGAGCTGATTATTTAATAGTAACAATGATTTAGGCAATAATTTCTATATTTTGTATTTTACTATTTCAATC  
GTACCTGTCATCGTATTGGCATCTGACTTATTTAAAGAAAGGATTTCAAAAGGTGTCTATTAATTGTATTGGCTATTATCGCT  
TTAGTATTATGCACTTTGTATCTGCAATCTCTGGTTTGTTCAGCCATATCTATTTTAGGTAGAAAAAATAGTAGCTGCA  
CGAGTACTACCATTATCAAAAAGTAAAGGGAACGCAATCAAGCATCACATAAAGACACGTGTAAAAAGGAACTTGATAGT  
CAAGACATGATGGAACACTCTGAGGTTAAAAATCCCACGACTAAAAACCTTGAAGGATTTAACGAAGAAATACATAAAGATGAA  
GCTACAATAAGTTGTCAAGTGATAACACGGAACCGCCTATTGAATCAAAAGACCATGTCTCGAAAAAGATTGATGACAACT  
AATCGAGAGACTTAAAAAATAATATTCAACATAAAGAACTTTTAAACGACATTTAAACGCATTGGCAATCACTAATGGTAGTG  
CGTTTAACTTAACTTAAATATCTGAATATTTTGTAAATGGAGCTACCTTTGTTGTTACTATTCAATGAAGAGGAGTAAATG  
TAATTAAGGAAAGAAATTTGAGGAGTGATCTTTATGACAAACAACAAGTAGCATTAGTAAGTGGCGGAGCACAAGGGATTGG  
TTTTAAAAATTGCAAGCTTTAGTGGAAGATGGTTTCAAGTAGCAGTGTGTTGATTTCAATGAAGAAGGGGCAAAAGCAGCTGC  
ACTTAAATCACTTAAATGATGGTACAAAAGCTATTGCTATCAAGAGATGTATCAAAACCGTGATGATGATTTTAAACGCAATAAGA  
CAAAGTCCGCGCAATTTGGCGATTTCCATGTCATGGTTAAACATGCCGCGCTTGGACCAACAACACCAATCGATACAACTACT  
GAAGAACAGTTTAAACAGTATATGGCGTGAACGTTGCAGGTGTGCTATGGGGTATTCAAGCCGCACATGAACAATTTAAAAA  
TTCAATCATGGCGGTAATTTATCAATGCAACATCTCAAGCAGGCGTTGAGGGTAACCCAGGCTTGTCTTTATATTGTCAGTACA  
AAATTCGAGTGGCGGTTTAAACAAGTAGCCGCAACAAGATTAGCGCTGAAGGTATTACTGTGAATGCATTTCGACCTGGT  
ATCGTTCAAAACACCAATGATGGAAGTATCGCAGTGGCAACAGCCGAAGAAGCAGGTAAACCTGAAGCATGGGGTTGGGAACAA  
TTTACAAGTCAGATTGCTTTGGGCAAGTTTCTCAACAGAGAAGATGTTTCAATGTAGTGAGCTTCTTACTGTTGTAAGACTCT  
GATTCAATATTCTGGACAACAATTTGTAGATGGTGGTATGAGATTCCGTTAATAATCATCCACTCAATGATAAATAAATCTCT  
ATTGTTAAGTTTAAATCACTTAGCAGTAAGGATTTTTTAGTGCATTTAGAAGGGAGTGTATTGGTAGAAAAATTAATAAGCGAAGT  
TCTTAAGTGAGTTATGATGTACAGTCTAATGCATCAGTTGAAAGCATTATTAGTATTAAACACACCCCAAGATATTATAAAACAT  
CACAAAACCACTATCTAATTTATCTCAATAAAAATTCACAAAGTTATCTCATTTTATTTTATAAATAAAAAATATCGATA  
AAAAGCTTACAATACCTTATGTTTTTATGATATATTTTTAATGTATAAATGAGGTGGAAGATTTGGAAAGAGTTTGTATAACTG  
GTGGGCTGTTTTTATGGGTCGCATTTAGTAGATGATTTACAACAAGATTATGATGTTTATGTTCTAGATAACTATAGAACAG  
GTAACGAGAAAAATATTAAAAGTTTGGCTGACGATCATGTGTTTGAATTAGATATTTCGTGAATATGATGCAGTTGAACAAATCA  
TGAAGACATATCAATTTGATTATGTTATTTCATTTAGCAGCATTAGTTAGTGTGCTGAGTGGTTGAGAAACCTATCTTATCTC  
AAGAAATAAACGTCGTAGCAACATTAAGATTGTTAGAAATCATTAATAATCATATAAAACGTTTTATCTTTGCTT  
CGTCAGCAGCTGTTTATGGTGATCTTCCGATTTGCTTAAAGCTGATCAATCATTAATCTTACCATTATCACCATTATGCAATAG  
ATAAATATTACGGCGAACCGGACGACATTAATTTGTTGCTTATATAACATACCAACAGCGTTGTTAAATTTTTTAATGTAT  
TTGGGCCAAGACAGGATCCTAAGTCACAAATATTCAGGTGTGATTTCAAAGATGTTTCGATTCAATTTGAGCATTAACAAGCATT  
CAATTTTTTGGTGACGACTGCAAACTAGAGATTTTGTATATGATATGATGTTGTTCAATCTGTACGCTTAATTATGGAACACA  
AAGATGCAATTGGACACGGTTATAACATTGGTACAGGCATTTTACTAATTTATTAGAGGTTTATCGTATTATTGGTGAATTA  
ATGAAAAATCAGTCAGCATGAATTTAAAGAAGCAGCAAGAGGAGATTAAGCATTTCTTAAGCATTTCTAAGCTTAAAGG  
CATTAGGATTTGTTCTTAAATATACAGTAGAAACAGGTTTAAAGGATTACTTTAATTTTGAAGTAGATAAATTGAAGAAGTTA  
CAGCTAAAGAAGTGGAAATGTCGTGAAATGACATTGAAGCTGTCCATAATAAAGGGTTATGCCTATCAAGAAAAATAGAC  
AAACTAAGAAGAGTGAGAAAAAGCTATTACCAATTTAAACGTCGATTTAATTTTAAAGCATTTTATTATTATTTTAACT  
TTACCGATTATGGTTATATTGCGCATTTGCTATCGTCATAGATTTCGCCAGGAAACCCCTATTTATAGTCAGGTTAGAGTTGGGAAG  
ATGGGTAAATTAATTTAAATATACAAATTTACGTTTCGATGTGCAAAAACCGCAGAGAAAAACGGTGCGCAATGGGCTGATAAAGAT  
GATGATCGTATAACAATATCGGGAAGTTTATTCGTAACACCGCATTTGATGAATACCACAATTAATGTTGTTAAAGGG  
GAAATGAGTTTTATTGGACACGCGCCGAACTGCGGAATTTGAGAATTTTAGTTTCAAGAGTATAGTTTTCGAGCAAGA  
TGCTTTGTTACACAGGGTTAACAGGACTTGCAGCAATTCAGGTGGATATGACTTAACACCGCAACAAAACTGAAATATGAC  
ATGAAATATATACATAAAGGTAGTTTAAATGATGGAACATATATATCAATTAGAACATTGATGGTTGTTATTACAGGGGAAGC

TCAAGGTAGTCTTAATTTACTTAAATAAGTTCAAATAAAAGTTATATTTTAAAGATTGTGACCAATTGTTACAGTATAACGAGGA  
ATCCCTTGAGACAGTATCAAATGGCATTAAGAAATATGTGCCATCATTGATTTGCATGGCTATAAACTACTATTCATCTGATGAG  
ATAGCCATGTTAAGAAATGAAAGTATAGCATTAAGGGGTTTGTAAACAGTTGAAAATATATATATTGTATTACTAAAGCAGACA  
ATGGTGGTGCACAAACACATCTCATTTCAACTCGCCAACCATTTTTGGCTACACAATGATGTTTATGTCTATTGTAGGCAATCATG  
GACCAATGATTGAACAACTAGATGCAAGAGTTAATGTAAATTTATTCGAACATTTAGTAGGTCCAATTGACTTTAAACAAGATA  
TTTTAGCTGTCAAAGTGTAGCACAGTTATTTCTCGAAAATTAACCTGATGTTATCCATTTACATTCTTCCAAAAGCTGGAACGG  
TCGGACGAATTGCGAAGTTCATTTTCGAAAATCGAAAGACACACGTATAGTTTTCACGACATGGATGGGCTTTTACAGAGGGTG  
TTAAACCAGCTAAAAATTTCTATATTTAGTTATCGAAAAATTAATGTCACCTATTACAGATAGCATTATTTGTGTTTCAGATT  
TCGATAAACAGTTAGCGTTAAAAATATCGATTTAATCGATTGAAATTAACCAACATACATAATGGTATTGCGATGTTCCCGCTG  
TTAAGCAAACGCTAAAAAGCCAATCACATAACAATATTGGCGAAGTAGTTGGAATGTTGCCATAATAACAAGATTTACAGATTA  
ATGCCCCGACAAAGCATCAATTTGTTATGATTGCAAGATTGCTTATCCAAAATTGGCCAAAAATCTAATCGCGGCAATAGAGA  
TATTGAAATTACATAACAGTAATCATGCGCATTTTACATTTATAGGCGATGGACCTACATTAAATGATTGTGACGAACAAGTTG  
TACAAGCTGGGTAGAAAATGATGTCACATTTTGGGCAATGTCATTAAATGCGAGTCATTTATTATCACAATACGATACGTTTA  
TTTTAATAAGTAAGCATGAAGGTTTGGCAATTAGCATTATAGAAGCTATGGCTACAGGTTTGCCTGTTATAGCCAGTCATGTTG  
GCGGTATTTAGAAATTAGTAGCTGATAATGTTATATGTATGATGAACAACCAACCGAACTATTGCTAAAGTCTGGAAGAAAT  
ATTTAATAGACAGTGAATTACATCAAAATGACTAATCAATCTAGAAAACGTTATTTAGAAATGTTTACTGAGGAGAAAATGATTA  
AAGAAGTGAAGACGTTTATAATGGAATAACAACAATAGTAAATTACTAACATTGTTACTTATCGGTTTAGCGGTTTTATT  
CAGCAATCTTCGGTTATTGCGGGTGTGAATGTTTCTATAGCTGACTTTATCACATTACTAATATTAGTTTATTTACTGTTTTCT  
GCTAACCATTTATTAAGGCAAAATCATTTTACAGTTTATAGGCGATGGACCTACATTAAATGATTGTATACATATCGTATTACGTTTGT  
CTATTTTGTGATGATTGATATTTATTACGGTTAAGGAAGTTCTTGCATCTACAGTTAAATATGCATTTGTAGTCATTTATTTT  
TATTTAGGATGATCTCTTTAAGTTAGGTAATAGCAAAAACGTTATGATGCTTACCTTATATTATAAGCAGTGTGACTATAGGT  
CTATTTTGTATTATAGCTGGTTTGAACAAGTCCCTTTACTAATGAATTTGTTATATTTTGTGATAATACGTTCAAAGGATTA  
ATGAATGACCCTAACTATTTGCGGATGACACAGATTATTACATTGGTACTTGCTTACAAGTATATTCTAATATCATATTCAAG  
GTCCTTGCATGTGGTATTTTGTCTATGGTCTTTAACTACAACGGGGTCTAAGACTGCGTTTATCATATTAACTCGTCTTAGCCATT  
TATTTCTTTTAAAAAAGTTATTTAGTAGAAAATGCGGTAAAGTGTGTGAGTATGTCAGTGATTATGCTGATATTACTTTGTTTT  
ACCTTTTATAATTAACACTACTATTTTATCCAAATTAAGCAACCTTGATGCGCTTACCCTCATTAGATCGAATGGCGCTATTTT  
GAAGAGGGCTTTGCATCATTAAATGATAGTGGGTCTGAGCGAAGTGTGTATGGATAAATGCCATTTTCAGTAATTAATATACA  
CTAGGTTTGGTGTGCGGATAGTGGATTATGTACATATTGGCTCGCAAAATTAATGGTATTTTACTTGTGCGCCATAATACATAT  
TTGAGATCTTTGCGGAATGGGGCATTTTATTCGGTGCATTTTATTCATATTATTTACTGCTTTTATTTGCTTTTGAATTTAGTA  
TTTAACTATTTCTGGGAAAAATGTAACAGCAATTGTTGTAATGTTGACGATGCTGATTACTTTTTAACAGTATCATTTAATAAC  
TCAAGATATGTCGCTTTTATTTTAGGAATATCGTCTTTATTTGTTCAATATGAAAAGATGGAAGGGATCGTAATGAAGAGTGA  
TTCATAAAAAGAAAATATTATTTATCAAGGGCTATACCAATTGATTAGAACGATGACACCACTGATTACAATACCCATTATTTT  
ACGTCTTTGGTCCGAGTGGTGGTATTTGTTTTCATTTTCTTTCAATATCGTGCAATACTTTTGTGATGATTGCAAGTGTGG  
CGTTCAAGTTATATTTTAAAGAGTTATCGCGAAGTCCGTTAAGCACAACGGCAATTGTCACAGCAGTTTGGGATATCTTTGT  
CAGTAAATATTTTACGCTTAAACAGTTTTCGCGATGTATATGGTCGTAATTACTATATTTATTTGATGATTACTATCTTATTT  
CCTACATAAGGAATCATATTATAGGTGACGACATCGATTTTATGCTGGAATGAAAAGTTTAAATTTCCCTAG  
CCTCAGTAATATTTGTTGCGTCTGGTATTGTATTAAGTGTAGTTGTTATTTTGTCAAAGATCAATCAGATTTATCATTTGTATGT  
ATTTACTATTGCTATTTGTGACGGTATTAACCAATTACCTTTGTTTATCTATTTTAAACGATACATATTAGCTTTGTTTGGTTAA  
TTGGATACACGCTCTGGCAATTGTTTCGTCATTAGCATTTATTTACCAATGGACAGCTCAACTTATATACATAGTATTTCT  
TTGCGTTGTTCTTGGTTTAGTAGGTACATACCAACAAGTTGGTATCTTTTCTAACGCATTTAATATTTTAAACGGTGCATCAT  
AATGATTAATACATTTGATCTTGTAAATGATTCCGCGTATTACCAAAATGTCTATCCAGCAATCACATAGTTTAACTAAAACGTT  
AGCTAATAATATGAATATTCAATTGATATTAACAATACCTATGGTCTTTGGTTTAAATGCAATTATGCCATCATTTTATTTATG  
GTTCTTTGGTGAGGAATTTCGCATCAACTGTCCCATTTGATGCAACCTTTAGCGATCTTGATTAATATCATTTCTTTAAATATGTT  
GATAAGCAGGCAATATTTATTAATAGTGAATAAAATAAGATTATATAATGCGTCAATTACTATTGGTGCAGTGATAAACCTAGT  
ATTATGTATTATTTTATATATTTTATGGAATTTACGGTGTGCTATTGCGCGTTTAAATACAGAGTTTCTTGTGCTCATTTG  
GCGATTATTTGATATTTACTAAAATCAATGTGAAGTTGAATATTTGTAAGTACGATTCAATGTGTCATTGCTGCTGCTATTGTT  
TATTGTGCTTGGTGTGGTCAATCATTTATTTGCCCCCTACAATGTACGCTACGCTGCTATTAATTGCGATTGGTATAGTAGTTTA  
TCCTTTTATTAATGATGACTATGAAAAATCAATACGTATGGCAAAATATTGAGGCATCTTCGACATAAAACAATTTAAGTACCGGT  
AATGCTATACTTTAGAAAATTAAGATTAGAAGAAAGGCAATTTCTTATGAAAAATGGAAGTTGCTTTTTTAAATTCCTTT  
AAAAGCGGAAACAAAAGCAGTTAAATGCCTTTTTCGATTCAATATTAATATTAATATCAATTTTAAATTTTAAATTTATAT  
AATTGGATATAACAATAAATAAATATTATGCAAAACACACCCAAAATTAATTTATATAAAAGTATATTCAATAAAAGGAGGAA  
TATACTTATGGCATTAAATACCAAAATTTACCATATGCATATGATGCATTGGAACCATATATAGATCAAGAACAATGGAGTT  
TCATCAGCAACAACATCAAAATACGTGACGAAATTAACGCAACAGTTGAAGGAACAGAGTTGAAGCAATCAATCACTACG  
GGATATGATTGCTAACTTAGACAAGGTACCGGAAGCGATGGGCTACCGAGCTCGAATTCGTAATCATGTATAGCTGTTTCTG  
TG

>168153\_3, SEQ ID NO:52

GTGGAAGATTGGAAGAGTTTTGATAACTGGTGGGGCTGGTTTTATTGGGTGCGATTTAGTAGATGATTACAACAAGATTAT  
GATGTTTATGTTCTAGATAACTATAGAACAGGTAAACGAGAAAATATTAAGTTTGGCTGACGATCATGTGTTGAATTAGAT  
ATTCGTGAATATGATGCAGTTGAACAAATCATGAAGACATATCAATTTGATTATGTTATTCATTTAGCAGCATAGTTAGTGT  
GCTGAGTCGGTTGAGAAACCTATCTTATCTCAAGAAATAAAGCTGATAGCAACATTAAGATTGTTAGAAATCATTAAGAAATAT  
AATAATCATATAAAACGTTTTATCTTTGCTTCGTCAGCAGCTGTTTATGGTGTCTTCTGATTGGCTAAAAGTGATCAATCA  
TTAATCTTACCATTTATCACCATATGAATAGATAAATATTACGGCGAACGGACGACATTAAATTTATGTTTCGTTATATAACATA  
CCAACAGCGGTTGTTAAATTTTTAATGATTTTGGGCCAAGCAGGATCTTAAGTCACAATATTCAGGTGTGATTCAAGATG  
TTCGATTCAATTTAGCATAACAAGCCATTACATTTTGGTGGAGGATGCAAACTAGAGATTTTGTATATGATGATGTT  
GTTCAATCTGTACGCTTAATTATGGAACACAAAGATGCAATTGGACAGGTTAATCAATTGGTACAGGCACTTTTACTAATTTA  
TTAGAGGTTTATCGTATTATTGGTGAATTATATGGAATAATCAGTCGAGCATGAATTTAAAGAAGCAGGAAAGGAGATATTAAG  
CATCTTATGCGATATTTCTAACTTAAAGGCATTAGGATTTGTTTCTTAAATATACAGTAGAAACAGGTTTAAAGGATTACTTT  
AATTTTGAAGTAGATAATATTGAAGAAGTTACAGCTAAAGAAGTGGAATGTCGTGA

>168153\_3, SEQ ID NO:53

VEDLERVLITGGAGFIGSHLVDDLQDDYDVYVLDNRYTGKRENIKSLADDDHVFELDIREDYDAVEQIMKTYQFDYVIHLAALVSV  
AESVEKPILSQEIINVVATLRLLEI IKYNNHIKRFIFASSAAVYGDLPDLPKSDQSLILPLSPYAI DKYYGERTTLNYSLYNI  
PTAVVKFFNVFGPRQDPKSSQYSGVISKMFDSFEHNKPTFFDGLQTRDFVYVYDVVQSVRLIMEHKDAIGHGYNIGTGTFNLT  
LEVYRIIGELYGKSVEHEFKEARKGDIKHSYADISNLKALGFVPKYTVETGLKDYFNFEVDNIEEVTAKEVEMS

>168153\_2, SEQ ID NO:54

ATGGTTATATTTCGCCATTGCTATCGTCATAGATTTCGCCAGGAAACCTATTTATAGTCAGGTTAGAGTTGGGAAGATGGGTAAA  
TTAATTTAAATATACAAATTACGTTTCGATGTGCAAAAACGCAGAGAAAAACGGTGCAGCAATGGGCTGATAAAGATGATGATCGT  
ATAACAAATGTTCGGGAAGTTTATTCGTAAAACACGCATTGATGAATTACCACAATAATTAATGTTGTTAAAGGGGAAATGAGT  
TTTATTGGACCACGCCCGGAACGTCGGAATTTGTAGAATTATTTAGTTTCAGAAGTGATAGGTTTCGAGCAAAGATGTCTTGT  
ACACCGGGTTAACAGGACTTGCAGCAAAATTCAGGTGGATATGACTTAACACCCGCAACAAAACTGAAATATGACATGAAATAT  
ATACATAAAGGTAGTTTAAATGATGGAACATATATATCAATTAGAACATTGATGTTGTTATTACAGGGGAAGGCTCAAGGTAG

>168153\_2, SEQ ID NO:55

LDKLEEVRSYPIKRAIDLILSIVLLFLTLPLIMVIFAIAIVIDSPGNPIYSQVRVGKMGKLIKIKYKLRSCKNAEKNGA  
QWADKDDDRITNVGKFIKTRIDELPQLINVVKGEMSFIGRPERPEFVELFSSEVIGFEQRCLVTPGLTGLAQIQGGYD  
LTPQQKLKYDMKYIHKGSLMELYSISIRTLMVVITGEGSR

>168153\_1, SEQ ID NO:56

ATGATTGAACAACCTAGATGCAAGAGTTAATGTAATTATTATCGAACATTTAGTAGGTCCAATTGACTTTAAACAAGATATTTTA  
GCTGTCAAAGTGTTAGCACAGTTATTCTCGAAAATTAAACCTGATGTTATCCATTTACATTTCTTCAAAGCTGGAACGGTCCGA  
CGAATTGCGAAGTTTCATTTGCGAAATCGAAAGACACACGATAGTTTACTGCACATGGATGGGCTTTTACAGAGGGTGTAA  
CCAGCTAAAAAATTTCTATATTTAGTTATCGAAAAATTAATGTCACCTATTACAGATAGCATTATTTGTGTTTCAGATTTTCGAT  
AAACAGTTAGCGTTAAAAATATCGATTTAATCGATTGAAATTAACCACAATACATAATGGTATTGCAGATGTTCCCGCTGTTAAG  
CAACCGCTAAAAAGCCAAATACATAACAATATGGCGAAGTAGTTGGAATGTTGCCCTAATAACAAGATTACAGATTAATGCC  
CCGACAAAGCATCAATTTGTTATGATTGCAAGATTTGCTTATCCAAAATGCCACAAAATCTAATCGCGGCAATAGAGATATTG  
AAATTACATAACAGTAATCATGCGCATTTTACATTTATAGGCGATGGACCTACATTAATGATTGTCAGCAACAAGTTGTACAA  
GCTGGGTTAGAAAAATGATGTCACATTTTGGGCAATGTCATTAATGCGAGTCATTTATTATCACAATACGATACGTTTATTTTA  
ATAAGTAAGCATGAAGGTTTGCCAATTAGCATTTATAGAAGCTATGGCTACAGGTTTGCCCTGTTATAGCCAGTCATGTTGGCGGT  
ATTTTCAGAAATTAGTAGCTGATAATGGTATATGTATGATGAACAACCAACCCGAACTATTGCTAAAGTCCTGAAAAATATTTA  
ATAGACAGTGATTACATCAAAATGAGTAATCAATCTAGAAAACGTTATTTAGAATGTTTACTGAGGAGAAAATGATTAAAGAA  
GTGGAAGACGTTTATAATGGAATCAACACAATAG

>168153\_1, SEQ ID NO:57

LKIIYICITKADNGGAQTHLIQLANHFVHNDVVIVGNHGMIEQLDARVNVIIIEHLVGPIDFKQDILAVKVLALQFSK  
IKPDVILHLSSKAGTVGRIAKFISKSKDTRIVFTAHGWAFTEGVKPAKKFLYLVIEKLSLITDSIIICVSDFDKQLALKY  
RFNRLKLTTHNGIADVPAVKQTLKSQSHNNIGEVVGMPLPNKQDLQINAPTKHQFVMIARFAYPKLPQNLIAAIEILKLH  
NSNHAHFTFIDGPTLNDCCQQVQAGLENDVTFGLNVINASHLLSQYDTFILISKHEGLPISIIEMATGLPVIASHVG  
GISELVADNGICMMNNQPETIAKVLEKYLIDSDYIKMSNQRKRYLECFTEEKMIKEVEDVYNGKSTQ

>168339\_1, SEQ ID NO:58, (ORF overlaps the 3' end of 168153\_1 by 20 nucleotides)

ATGGAATCAACACAATAGTAAATTACTAACATTGTTACTTATCGGTTTAGCGGTTTTATTACGCAATCTTCGGTTATTGCC  
GGTGTGAATGTTTCTATAGCTGACTTTATCATTACTAATATTAGTTTATTTACTGTTTTTCGCTAACCAATTTATTAAAGGCA  
AATCATTTTTTACAGTTTTTCATTATTTTGTATACATATCGTATGATTATTACGCTTTGTTTGTCTATTTTGTATGATTGATA  
TTTATTACGGTTAAGGAAGTTCTTGCATCTACAGTTAAATATGCATTTGTAGTCATTTATTTCTATTATTAGGATGATCATCTT  
AAGTTAGGTAATAGCAAAAAAGTGATCGTTACCTCTTATATTATAAGCAGTGTGACTATAGGTCTATTTTGTATTATAGCTGGT  
TTGAACAAGTCCCCCTTACTAATGAAATGTTATATTTTGTATGAAATACGTTCAAAGGATTAATGAATGACCCTAAGTATTTT  
GCCATGACACAGATTTTACATTGGTACTTGCCTTACAAGTATATTCAATATTACATATTCAAGGTCCTTGCATGCGTATTTT  
CTATGGTCTTTAACTACAACGGGGTCTAAGACTGCGTTTATCATATTAATCGTCTTAGCCATTTATTTCTTTATTAAGGTTA  
TTTAGTAGAAATGCGGTAAGTGTTGTGAGTATGTCAGTGATTATGCTGATATTACTTTGTTTACCTTTTATAATATCAACTAC  
TATTTATTCCAATTAAGCGACCTTGATGCCTTACCGTTCATTAGATCGAATGGCGTCTATTTTGAAGAGGGCTTTGCATCATT  
AATGATAGTGGGTCTGAGCGAAGTGTTGTATGGATAAATGCCATTTTCAGTAATTAATATACACTAGGTTTGGTGTGCGGATTA  
GTGATTATGTACATATTGGCTCGCAATTAATGGTATTTTACTTGTGTTGCCATAATACATATTTGCAGATCTTTGCGGAATGG  
GGCATTTTATCGGTGCATTTATTTATCATATTTGCTTTTATTACTGTTTGAATTTATAGATTTAACATTTCTGGGAAAAAT  
GTAACCGCAATTTGTTGTAATGTTGACGATGCTGATTTACTTTTACAGTATCATTTAATAACTCAAGATATGTCGCTTTTATT  
TTAGGAATTATCGTCTTTATTGTTCAATATGAAAAGATGGAAGGGATCGTAATGAAGAGTGA

>168339\_1, SEQ ID NO:59

MENQHNSKLLTLLLLIGLAVFIQSSVIAGVNVSIADFITLLILVYLLFFANHLLKANHFLOFFIILYTYRMIITLCLLFFDDLI  
FITVKEVLASTVKYAFVVIYFYLGMIIFKLGNSSKVVITSYIISSVTIGLFCIIAGLNKSPLLMKLLYFDEIRSKGLMNDPNYF  
AMTQIIITLVLAYKYIHNYIFKVLACGILLWSLTTTGSKTAFIILVLAIYFFIKKLFSRNAVSVVSMVIMILLGCTTFYNINY  
YLFQLSDLDALPSLDRMASIFEEGFASLNDSGSERSVVWINAISVIKYTLGFGVGLVDYVHIGSQINGILLVAHNTYLQIFAEW  
GILFGALFIIFMLYLLFELFRFNI SGKNVTAIVVMLTMLIYFLTVSFNNSRYVAFILGIIVFIVQYEMERDRNEE

>168339\_2, SEQ ID NO:60, (ORF overlaps the 3' end of 168339\_1 by 35 nucleotides)

ATGAAAAGATGGAAGGGATCGTAATGAAGAGTGATTCACTAAAAGAAAATATTATTATCAAGGGCTATACCAATTGATTAGA  
ACGATGACACCCTGATTACAATACCCATTATTTACGTCATTTGGTCCCAGTGGTGTGGGTATTGTTTCAATTTCTTTCAAT  
ATCGTGCAACTATTTTGTATGATTGCAAGTGTGGCGTTACGTTATATTTTAAATAGAGTTATCGCGAAGTCGGTTAACGCAAA  
CGGCAATTGTACAGCAGTTTGGGATATCTTTGTGCTAGTAAATATTTTATAGCGTTAACAGTTTGGCGATGATATGGTCGTA  
ATTACTATATTTATGATGATTACTATCTTATTTTCTACTACAAGGAATCTATATTATAGGTGCAGCACTCGATATTTTCATGG  
TTTTATGCTGGAACGAAAAGTTTAAATTCCTAGCCTCAGTAATATTGTTGCGTCTGGTATTGTTAAGTGATAGTTGTTATT  
TTTGTCAAAGATCAATCAGATTTATCATTTGATGTTTACTATTTGCTATTGTTGACGGTATTAACCAATTTACCTTTGTTTATC  
TATTTAAACGATACATTAGCTTTGTTTTCGGTTAATTGGATACAGCTCTGGCAATGTTTTCGTTTCGTCATTAGCATACTTATTA  
CCAAATGGACAGCTCAACTATATACCTAGTATTTTTCGCTTGTCTTGGTTTGTAGGTACATACCAACAAGTTGGTATCTTT  
TCTAACGCAATTAATATTTTAAACGGTCGCAATCAATGATTAATACATATTTGATCTTGTAATGATTCGGCTATTACCAAAATG  
TCTATCCAGCAATCACATAGTTTAACTAAAACGTTAGCTAATAATATGAATATTTCAATTGATATTAACAATACCTATGGTCTTT  
GGTTAATTGCAATTTATGCCATCATTTTATTTATGGTTCTTTGGTGAGGAATTCGCATCAACTGTCCCATTTGATGACCATTTTA  
GCGATACTTGTATTATCATTCCTTTAAATATGTTGATAAGCAGCAATATTTATTAATAGTGAATAAAAAAAGATTATATAAT

```
GCGTCAATTACTATTGGTGCAGTGATAAACCTAGTATTATGTATTATTTTGATATATTTTATGGAATTTACGGTGCTGCTATT
GCGCGTTTAATTACAGAGTTTTTCTTGCTCATTGGCGATTATTTGATATTACTAAATCAATGTGAAGTTGAATATTGTAAGT
ACGATTCAATGTGTGCTGCTGCTGTTATGATGTTTATTGTGCTGGTGTGGTCAATCATTATTTGCCCCCTACAATGTACGCT
ACGCTGCTATTAAATTGCGATTGGTATAGTAGTTATCTTTTATTAATGATGACTATGAAAAATCAATACGTATGGCAAATATTG
AGGCATCTTCGACATAAAACAATTTAA
```

>168339\_2, SEQ ID NO:61

```
MKSDSLKENIIYQGLYLIRTMPLITIPISRAFGPSGVGIVSFSFNIVQYFLMIASVGVQLYFNRVIAKSVNDRQLS
QQFWDIFVSKLFLALTVFAMYMVITIFIDDYLIFFLLQGIYIIIGAALDISWFYAGTEKFKIPSLSNIVASGIVLSVVVI
FVKDQSDLSLYVFTIAIVTVLNQLPLFIYLYKRYISFVSVNWIHVWQLFRSSLAYLLPNGQLNLNLYTSISCVVLGLVGTYYQ
VGIFSNAFNILTVAIIMINTFDLVMIPRITKMSIQQSHSLTKTLANNMNIQLILTIPMVFGLIAIMPSFYLWFFGEEFAS
TVPLMTILAILVLIIPLNMLISRQYLLIVNKIRLYNASITIGAVINLVLCIILYFYGIYGAAIARLITEFFLLIWRFD
ITKINVKLNIVSTIQCVIAAVMMFIVLGVVNHYPPTMYATLLLIAIGIVVYLLMMTMKNQYVWQILRHLRHKTI
```

**[0062]** Nucleic acid molecules of the present invention may be in the form of RNA, such as mRNA, or in the form of DNA, including, for instance, DNA and genomic DNA obtained by cloning or produced synthetically. The DNA may be double-stranded or single-stranded. Single-stranded DNA or RNA may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand.

**[0063]** The present invention further encompasses nucleic acid molecules of the present invention that are chemically synthesized, or reproduced as peptide nucleic acids (PNA), or according to other methods known in the art. The use of PNAs would serve as the preferred form if the nucleic acid molecules of the invention are incorporated onto a solid support, or gene chip. For the purposes of the present invention, a peptide nucleic acid (PNA) is a polyamide type of DNA analog and the monomeric units for adenine, guanine, thymine and cytosine are available commercially (Perceptive Biosystems). Certain components of DNA, such as phosphorus, phosphorus oxides, or deoxyribose derivatives, are not present in PNAs. As disclosed by P. E. Nielsen, M. Egholm, R. H. Berg and O. Buchardt, *Science* 254, 1497 (1991); and M. Egholm, O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden, and P. E. Nielsen, *Nature* 365, 666 (1993),

**[0064]** PNAs bind specifically and tightly to complementary DNA strands and are not degraded by nucleases. In fact, a PNA binds more strongly to DNA than does DNA itself. This is probably because there is no electrostatic repulsion between the two strands, and also the polyamide backbone is more flexible. Because of this, PNA/DNA duplexes bind under a wider range of stringency conditions than DNA/DNA duplexes, making it easier to perform multiplex hybridization. Smaller probes can be used than with DNA due to the strong binding. In addition, it is more likely that single base mismatches can be determined with PNA/DNA hybridization because a single mismatch in a PNA/DNA 15-mer lowers the melting point ( $T^{sub.m}$ ) by 8°-20° C, vs. 4°-16° C for the DNA/DNA 15-mer duplex. Also, the absence of charge groups in PNA means that hybridization can be done at low ionic strengths and reduce possible interference by salt during the analysis.

[0065] By "isolated" polynucleotide sequence is intended a nucleic acid molecule, DNA or RNA, which has been removed from its native environment. This includes segments of DNA comprising the *S. aureus* polynucleotides of the present invention isolated from the native chromosome. These fragments include both isolated fragments consisting only of *S. aureus* DNA and fragments comprising heterologous sequences such as vector sequences or other foreign DNA. For example, recombinant DNA molecules contained in a vector are considered isolated for the purposes of the present invention which may be partially or substantially purified. Further examples of isolated DNA molecules include recombinant DNA molecules introduced and maintained in heterologous host cells or purified (partially or substantially) DNA molecules in solution. Isolated RNA molecules include *in vivo* or *in vitro* RNA transcripts of the DNA molecules of the present invention. Isolated nucleic acid molecules according to the present invention further include such molecules produced synthetically which may be partially or substantially purified the excluded RNA or heterologous DNA. Isolated nucleic acid molecules at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 96%, 98%, 99%, or 100% pure relative to hereologous (*Staphylococcus* or other) (DNA or RNA) or relative to all materials and compounds other than the carrier solution. The term "isolated" does not refer to genomic or cDNA libraries, whole cell mRNA preparations, genomic DNA digests (including those gel separated by electrophoresis), whole chromosome or sheared whole cell genomic DNA preparations or other compositions where the art demonstrates no distinguishing features of the polynucleotides sequences of the present invention.

[0066] In addition, isolated nucleic acid molecules of the invention include DNA molecules which comprise a sequence substantially different from those described above but which, due to the degeneracy of the genetic code, still encode a *S. aureus* polypeptides and peptides of the present invention (e.g., polypeptides of Table 1). That is, all possible DNA sequences that encode the *S. aureus* polypeptides of the present invention. This includes the genetic code and species-specific codon preferences known in the art. Thus, it would be routine for one skilled in the art to generate the degenerate variants described above, for instance, to optimize codon expression for a particular host (e.g., change codons in the bacteria mRNA to those preferred by a mammalian or other bacterial host such as *E. coli*).

[0067] The invention further provides isolated nucleic acid molecules having the nucleotide sequence shown in Table 1 or a nucleic acid molecule having a sequence complementary to one of the above sequences. Such isolated molecules, particularly DNA molecules, are useful as probes for gene mapping and for identifying *S. aureus* in a biological

sample, for instance, by PCR or hybridization analysis (e.g., including but not limited to, Northern blot analysis). In specific embodiments, the polynucleotides of the present invention are less than 300kb, 200kb, 100kb, 50kb, 10kb, 7.5kb, 5kb, 2.5kb, and 1kb. In another embodiment, the polynucleotides comprising the coding sequence for polypeptides of the present invention do not contain genomic flanking gene sequences or contain only genomic flanking gene sequences having regulatory control sequences for the said polynucleotides.

**[0068]** The present invention is further directed to nucleic acid molecules encoding portions or fragments of the nucleotide sequences described herein. Uses for the polynucleotide fragments of the present invention include probes, primers, molecular weight markers and for expressing the polypeptide fragments of the present invention. Fragments include portions of the nucleotide sequences of Table 1, at least 10 contiguous nucleotides in length selected from any two integers, one of which representing a 5' nucleotide position and a second of which representing a 3' nucleotide position, where the first nucleotide for each nucleotide sequence in Table 1 is position 1. That is, every combination of a 5' and 3' nucleotide position that a fragment at least 10 contiguous nucleotides in length could occupy is included in the invention as an individual species. "At least" means a fragment may be 10 contiguous nucleotide bases in length or any integer between 10 and the length of an entire nucleotide sequence minus 1. Therefore, included in the invention are contiguous fragments specified by any 5' and 3' nucleotide base positions of a nucleotide sequences of Table 1 wherein the contiguous fragment is any integer between 10 and the length of an entire nucleotide sequence minus 1.

**[0069]** The polynucleotide fragment specified by 5' and 3' positions can be immediately envisaged using the clone description and are therefore not individually listed solely for the purpose of not unnecessarily lengthening the specifications.

**[0070]** Although it is particularly pointed out that each of the above described species may be included in or excluded from the present invention. The above species of polynucleotides fragments of the present invention may alternatively be described by the formula "a to b"; where "a" equals the 5' nucleotide position and "b" equals 3' nucleotide position of the polynucleotide fragment, where "a" equals an integer between 1 and the number of nucleotides of the polynucleotide sequence of the present invention minus 10, where "b" equals an integer between 10 and the number of nucleotides of the polynucleotide sequence of the present invention; and where "a" is an integer smaller than "b" by at least 10.

**[0071]** Again, it is particularly pointed out that each species of the above formula may be specifically included in, or excluded from, the present invention. Further, the invention



includes polynucleotides comprising sub-genuses of fragments specified by size, in nucleotides, rather than by nucleotide positions. The invention includes any fragment size, in contiguous nucleotides, selected from integers between 10 and the length of an entire nucleotide sequence minus 1. Preferred size of contiguous nucleotide fragments include at least 20 nucleotides, at least 30 nucleotides, at least 40 nucleotides, at least 50 nucleotides, at least 60 nucleotides, at least 70 nucleotides, at least 80 nucleotides, at least 90 nucleotides, at least 100 nucleotides, at least 125 nucleotides, at least 150 nucleotides, at least 175 nucleotides, at least 200 nucleotides, at least 250 nucleotides, at least 300 nucleotides, at least 350 nucleotides, at least 400 nucleotides, at least 450 nucleotides, at least 500 nucleotides, at least 550 nucleotides, at least 600 nucleotides, at least 650 nucleotides, at least 700 nucleotides, at least 750 nucleotides, at least 800 nucleotides, at least 850 nucleotides, at least 900 nucleotides, at least 950 nucleotides, at least 1000 nucleotides, at least 1050 nucleotides, at least 1100 nucleotides, and at least 1150 nucleotides. Other preferred sizes of contiguous polynucleotide fragments, which may be useful as diagnostic probes and primers, include fragment sizes representing each integer between 50-300. Larger fragments are also useful according to the present invention corresponding to most, if not all, of the polynucleotide sequences of the sequence listing or Table 1. The preferred sizes are, of course, meant to exemplify not limit to present invention as all size fragments, representing any integer between 10 and the length of an entire nucleotide sequence minus 1 of the sequence listing or deposited clones, may be specifically included from the invention. Additional preferred nucleic acid fragment of the present invention include nucleic acid molecules encoding epitope-bearing portions of the polynucleotides (e.g., including but not limited to, nucleic acid molecules encoding epitope-bearing portions of the polynucleotides which are shown in Table 4).

**[0072]** In another aspect, the invention provides an isolated nucleic acid molecule comprising a polynucleotide which hybridizes under stringent hybridization conditions to a portion of a polynucleotide in a nucleic acid molecules of the invention described above, for instance, nucleotide sequences of Table 1. By "stringent hybridization conditions" is intended overnight incubation at 42°C in a solution comprising: 50% formamide, 5x SSC (750 mM NaCl, 75 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 µg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C. Hybridizing polynucleotides are useful as diagnostic probes and primers as discussed above. Portions of a polynucleotide which hybridize to a nucleotide sequence in Table 1, which can be used as probes and primers, may be precisely specified by 5' and 3' base positions or by size in nucleotide bases as described

above or precisely excluded in the same manner. Preferred hybridizing polynucleotides of the present invention are those that, when labeled and used in a hybridization assay known in the art (e.g. Southern and Northern blot analysis), display the greatest signal strength with the polynucleotides of Table 1 regardless of other heterologous sequences present in equimolar amounts

**[0073]** The nucleic acid molecules of the present invention, which encode a *S. aureus* polypeptide, may include, but are not limited to, nucleic acid molecules encoding the full length *S. aureus* polypeptides of Table 1. Also included in the present invention are nucleic acids encoding the above full length sequences and further comprise additional sequences, such as those encoding an added secretory leader sequence, such as a pre-, or pro- or prepro-protein sequence. Further included in the present invention are nucleic acids encoding the above full length sequences and portions thereof and further comprise additional heterologous amino acid sequences encoded by nucleic acid sequences from a different source.

**[0074]** Also included in the present invention are nucleic acids encoding the above protein sequences together with additional, non-coding sequences, including for example, but not limited to non-coding 5' and 3' sequences. These sequences include transcribed, non-translated sequences that may play a role in transcription, and mRNA processing, for example, ribosome binding and stability of mRNA. Also included in the present invention are additional coding sequences which provide additional functionalities.

**[0075]** Thus, a nucleotide sequence encoding a polypeptide may be fused to a marker sequence, such as a sequence encoding a peptide which facilitates purification of the fused polypeptide. In certain preferred embodiments of this aspect of the invention, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. For instance, hexa-histidine provides for convenient purification of the fusion protein. See Gentz et al. (1989) Proc. Natl. Acad. Sci. 86:821-24. The "HA" tag is another peptide useful for purification which corresponds to an epitope derived from the influenza hemagglutinin protein. See Wilson et al. (1984) Cell 37:767. As discussed below, other such fusion proteins include the *S. aureus* fused to Fc at the N- or C-terminus.

**[0076] Variant and Mutant Polynucleotides**

**[0077]** The present invention further relates to variants of the nucleic acid molecules which encode portions, analogs or derivatives of a *S. aureus* polypeptides of Table 1, and variant polypeptides thereof including portions, analogs, and derivatives of the *S. aureus*

polypeptides. Variants may occur naturally, such as a natural allelic variant. By an "allelic variant" is intended one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. See, *e.g.*, B. Lewin, Genes IV (1990). Non-naturally occurring variants may be produced using art-known mutagenesis techniques.

**[0078]** Such nucleic acid variants include those produced by nucleotide substitutions, deletions, or additions. The substitutions, deletions, or additions may involve one or more nucleotides. The variants may be altered in coding regions, non-coding regions, or both. Alterations in the coding regions may produce conservative or non-conservative amino acid substitutions, deletions or additions. Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of a *S. aureus* protein of the present invention or portions thereof. Also preferred in this regard are conservative substitutions.

**[0079]** Such polypeptide variants include those produced by amino acid substitutions, deletions or additions. The substitutions, deletions, or additions may involve one or more residues. Alterations may produce conservative or non-conservative amino acid substitutions, deletions, or additions. Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of a *S. aureus* protein of the present invention or portions thereof. Also especially preferred in this regard are conservative substitutions.

**[0080]** The present invention also relates to recombinant vectors, which include the isolated nucleic acid molecules of the present invention, and to host cells containing the recombinant vectors, as well as to methods of making such vectors and host cells and for using them for production of *S. aureus* polypeptides or peptides by recombinant techniques.

**[0081]** The present application is directed to nucleic acid molecules at least 90%, 95%, 96%, 97%, 98% or 99% identical to a nucleic acid sequence shown in Table 1. The above nucleic acid sequences are included irrespective of whether they encode a polypeptide having *S. aureus* activity. This is because even where a particular nucleic acid molecule does not encode a polypeptide having *S. aureus* activity, one of skill in the art would still know how to use the nucleic acid molecule, for instance, as a hybridization probe or primer. Uses of the nucleic acid molecules of the present invention that do not encode a polypeptide having *S. aureus* activity include, *inter alia*, isolating an *S. aureus* gene or allelic variants thereof from a DNA library, and detecting *S. aureus* mRNA expression in biological or environmental samples, suspected of containing *S. aureus* by Northern Blot analysis or PCR.

**[0082]** For example, one such method involves assaying for the expression of a

polynucleotide encoding *S. aureus* polypeptides in a sample from an animal host (e.g., including, but not limited to, human, bovine, rabbit, porcine, murine, chicken, and/or avian species). The expression of polynucleotides can be assayed by detecting the nucleic acids of Table 1. An example of such a method involves the use of the polymerase chain reaction (PCR) to amplify and detect *Staphylococcus* nucleic acid sequences in a biological or environmental sample.

**[0083]** The present invention also relates to nucleic acid probes having all or part of a nucleotide sequence described in Table 1 which are capable of hybridizing under stringent conditions to *Staphylococcus* nucleic acids. The invention further relates to a method of detecting one or more *Staphylococcus* nucleic acids in a biological sample obtained from an animal, said one or more nucleic acids encoding *Staphylococcus* polypeptides, comprising: (a) contacting the sample with one or more of the above-described nucleic acid probes, under conditions such that hybridization occurs, and (b) detecting hybridization of said one or more probes to the *Staphylococcus* nucleic acid present in the biological sample.

**[0084]** The invention also includes a kit for analyzing samples for the presence of members of the *Staphylococcus* genus in a biological or environmental sample. In a general embodiment, the kit includes at least one polynucleotide probe containing a nucleotide sequence that will specifically hybridize with a *S. aureus* nucleic acid molecule of Table 1 and a suitable container. In a specific embodiment, the kit includes two polynucleotide probes defining an internal region of the *S. aureus* nucleic acid molecule of Table 1, where each probe has one strand containing a 31'-mer-end internal to the region. In a further embodiment, the probes may be useful as primers for polymerase chain reaction amplification.

**[0085]** The method(s) provided above may preferably be applied in a diagnostic method and/or kits in which *S. aureus* polynucleotides of Table 1 are attached to a solid support. In one exemplary method, the support may be a "gene chip" or a "biological chip" as described in US Patents 5,837,832, 5,874,219, and 5,856,174. Further, such a gene chip with *S. aureus* polynucleotides of Table 1 attached may be used to diagnose *S. aureus* infection in an animal host, preferably a human. The US Patents referenced above are incorporated herein by reference in their entirety.

**[0086]** The present invention is further directed to nucleic acid molecules having sequences at least 90%, 95%, 96%, 97%, 98% or 99% identical to the nucleic acid sequence shown in Table 1, which do, in fact, encode a polypeptide having *S. aureus* protein activity. By "a polypeptide having *S. aureus* activity" is intended polypeptides exhibiting activity similar, but not necessarily identical, to an activity of the *S. aureus* protein of the invention, as

measured in a particular biological assay suitable for measuring activity of the specified protein. The biological activity of some of the polypeptides of the presents invention are listed in Table 1, after the name of the closest homolog with similar activity. The biological activities were determined using methods known in the art for the particular biological activity listed. For the remaining polypeptides of Table 1, the assays known in the art to measure the activity of the polypeptides of Table 2, sharing a high degree of identity, may be used to measure the activity of the corresponding polypeptides of Table 1.

**[0087]** Of course, due to the degeneracy of the genetic code, one of ordinary skill in the art will immediately recognize that a large number of the nucleic acid molecules having a sequence at least 90%, 95%, 96%, 97%, 98%, or 99% identical to the nucleic acid sequences shown in Table 1 will encode a polypeptide having biological activity. In fact, since degenerate variants of these nucleotide sequences all encode the same polypeptide, this will be clear to the skilled artisan even without performing the above described comparison assay. It will be further recognized in the art that, for such nucleic acid molecules that are not degenerate variants, a reasonable number will also encode a polypeptide having biological activity. This is because the skilled artisan is fully aware of amino acid substitutions that are either less likely or not likely to significantly effect protein function (e.g., replacing one aliphatic amino acid with a second aliphatic amino acid), as further described below.

**[0088]** By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence of the present invention, it is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the *S. aureus* polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted, inserted, or substituted with another nucleotide. The query sequence may be an entire sequence shown in Table 1, the ORF (open reading frame), or any fragment specified as described herein.

**[0089]** Other methods of determining and defining whether any particular nucleic acid molecule or polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to a nucleotide sequence of the presence invention can be done by using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of

Brutlag et al. *See* Brutlag et al. (1990) *Comp. App. Biosci.* 6:237-245. In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by first converting U's to T's. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identity are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 0.05, Window Size=500 or the length of the subject nucleotide sequence, whichever is shorter.

TABLE 2. Closest matching sequence between the polypeptides of the present invention and sequences in GenSeq and GenBank databases

Sequence ID	Antigen Accession No.	Match Gene Name	High Score	Smallest Sum Probability P (N)
<i>GenSeq</i>				
HGS001	W34207	Streptomyces fabH homologue (frenolicin gene I pro...	285	3.50E-65
HGS001	W55808	Streptomyces roseofulvus frenolicin gene cluster p...	285	3.50E-65
HGS002	W20949	H. pylori cytoplasmic protein, 29zp10241orf7.	81	5.10E-12
HGS003	W48300	Staphylococcus aureus Fab I enoyl-ACP reductase.	1271	1.90E-170
HGS003	W40806	M. bovis InhA protein.	95	1.00E-29
HGS003	R23793	Stearoyl-ACP-desaturase (from clone pDES7).	157	1.60E-28
HGS003	R66290	M. tuberculosis inhA gene.	94	7.40E-28
HGS003	R66901	M. tuberculosis InhA.	94	7.40E-28
HGS003	R66292	Mycobacterium bovis InhA.	92	4.70E-19
HGS003	R63900	M. bovis InhA.	92	4.70E-19
HGS003	W16684	Lawsonia intracellularis enoyl-(acyl carrier prote...	114	1.80E-09
HGS003	W40805	M. tuberculosis InhA protein.	96	2.60E-09
HGS003	W40807	M. smegmatis InhA protein, mc2155 inhA-1.	101	9.70E-09
HGS004	W32287	Streptococcus pneumoniae MurA protein.	643	4.00E-89
HGS004	W26786	Streptococcus pneumoniae Mur A-1.	643	4.10E-89
HGS004	W27782	UDP-N-acetylglucosamine 1-carboxyvinyltransferase.	163	1.80E-15
HGS004	W27783	UDP-N-acetylglucosamine 1-carboxyvinyltransferase.	120	1.90E-12
HGS006	W36168	Staphylococcus aureus SP protein.	584	4.30E-78
HGS006	W37468	Staphylococcus aureus RNase P.	581	1.10E-77
HGS007M	W27798	Amino acid sequence of a replicative DNA heli case	5524	6e-83.2
HGS007M	R29636	pCTD ORF 1.	241.	7e-34.3
HGS008	W27814	A malonyl coenzymeA-acyl carrier protein transacyl...	365	4.70E-46
HGS008	W19629	Streptomyces venezuelae polyketide synthase.	96	2.30E-19
HGS008	W22602	Tylactone synthase ORF2 protein.	83	2.90E-18
HGS008	W22605	Tylactone synthase ORF5 protein.	95	8.90E-17
HGS008	R44431	eryA region polypeptide module #2.	88	2.30E-14
HGS008	R42452	Enzyme involved in eicosapentaenoic acid (EPA) syn...	94	5.30E-14

HGS008	R99462	Biosynthetic enzyme of icosapentaenoic acid synthase.	94	4.60E-13
HGS008	W37050	<i>S. putrefaciens</i> EPO biosynthesis gene cluster ORF6...	94	4.60E-13
HGS008	R44432	eryA region polypeptide module #3.	83	6.20E-13
HGS008	W22607	Platenolide synthase ORF2 protein.	80	2.20E-12
HGS014	W34454	Racillus subtilis teichoic acid polymerase.	597	2.70E-87
HGS014	W34455	Racillus subtilis teichoic acid polymerase.	597	3.10E-87
HGS014	W27744	Amino acid sequence of teichoic acid biosynthesis p...	425	2.50E-53
HGS016	W32287	Streptococcus pneumoniae MurA protein.	643	4.00E-89
HGS016	W26786	Streptococcus pneumoniae Mur A-1.	643	4.10E-89
HGS016	W27782	UDP-N-acetylglucosamine 1-carboxyvinyltransferase.	163	1.80E-15
HGS016	W27783	UDP-N-acetylglucosamine 1-carboxyvinyltransferase.	120	1.90E-12
HGS018	R95648	Thermostable DNA-ligase.	833	3.00E-205
HGS018	R81473	Thermus aquaticus DNA ligase protein.	428	2.00E-201
HGS018	R15299	Thermostable T. aquaticus ligase (I).	428	7.40E-199
HGS018	R15694	Thermostable T. aquaticus ligase (II).	428	4.80E-196
HGS019	P70096	Met-aminopeptidase.	143	2.90E-35
HGS019	R90027	Methionine aminopeptidase sequence.	138	1.60E-20
HGS022	R12401	Enantioselective amidase of Rhodococcus.	405	4.70E-102
HGS022	R25320	Enantioselective amidase.	405	4.70E-102
HGS022	W14159	Rhodococcus rhodochrous amidase.	352	6.10E-63
HGS022	W17820	Pseudomonas putida amidase.	208	1.20E-62
HGS022	R12400	Enantioselective amidase of Brevibacterium.	353	2.90E-62
HGS022	R24529	Enantioselective amidase.	353	2.90E-62
HGS022	W10882	Comamonas acidovorans derived amidase enzyme.	261	4.00E-61
HGS022	R60155	Comamonas testosteroni NI 1 amidase.	306	5.30E-47
HGS022	R42839	Urea amidolyase.	243	1.40E-31
HGS022	R44504	Urea amide lyase.	224	8.60E-30
HGS026	W29380	<i>S. pneumoniae</i> peptide releasing factor RF-1.	593	3.30E-142
HGS028	W29380	<i>S. pneumoniae</i> peptide releasing factor RF-1.	218	1.70E-49
HGS031	W20646	<i>H. pylori</i> cytoplasmic protein, O2cp11822orf26.	291	5.70E-47
HGS031	W20147	<i>H. pylori</i> cytoplasmic protein, 14574201.aa.	75	1.50E-08
HGS033	W20861	<i>H. pylori</i> cell envelope transporter protein, 12gel...	100	2.30E-18
HGS033	W20101	<i>H. pylori</i> transporter protein 11132778.aa.	100	6.10E-17



HGS033	W25671	hABC3 protein.	111	4.20E-15
HGS033	W46761	Amino acid sequence of human ATP binding cassette ...	111	4.20E-15
HGS033	W46771	Amino acid sequence of human ATP binding cassette ...	111	4.30E-15
HGS033	W42393	Bacillus thermoleovorans phosphatase (68FY5).	96	1.90E-13
HGS033	W34202	Streptomyces efflux pump protein (frenolicin gene ...	92	5.50E-12
HGS033	W55803	Streptomyces roseofulvus frenolicin gene cluster p...	92	5.50E-12
HGS033	W20224	H. pylori transporter protein, 22265691.aa.	88	7.40E-12
HGS033	W20668	H. pylori transporter protein Q3ee11215orf29.	88	8.90E-12
HGS036	W20640	H. pylori transporter protein, 02ce11022orf8.	264	2.20E-33
HGS036	W34202	Streptomyces efflux pump protein (frenolicin gene ...	184	1.30E-29
HGS036	W55803	Streptomyces roseofulvus frenolicin gene cluster p...	184	1.30E-29
HGS036	W20289	H. pylori transporter protein, 24218968.aa.	201	5.50E-21
HGS036	W20711	H. pylori transporter protein, 05cp11911orf41.	148	2.10E-19
HGS036	W20101	H. pylori transporter protein 11132778.aa.	164	3.50E-19
HGS036	W20861	H. pylori cell envelope transporter protein, 12gel...	164	4.20E-19
HGS036	W20492	H. pylori cell envelope transporter protein 433843...	148	1.60E-18
HGS036	W21019	H. pylori cell envelope transporter protein, hp5e1...	144	8.30E-16
HGS036	R71091	C. jejuni PEB1A antigen from ORF3.	136	7.90E-14
168153_3	W01619	Human uridine diphosphate galactose-4-epimerase.	128	9.80E-29
168153_3	W40383	S. glaucescens acbD protein.	105	1.10E-15
168153_3	R98529	dTDP-glucose dehydratase encoded by the acbB gene.	108	4.50E-15
168153_3	R80287	galE gene of S. lividans gal operon.	88	2.60E-13
168153_3	P70275	Sequence encoded by S.lividans gal operon galE gene.	86	5.10E-13
168153_3	R41529	S.lividans UDP-4-epimerase.	86	5.10E-13
168153_3	R32195	ADP-L-glycero-D-mannoheptose-6-epimerase protein.	82	3.40E-10
168153_2	W03997	Glucosyl IP-transferase (SpsB protein).	168	8.30E-36
168153_2	W32794	Sphingomonas genus microbe isolated SpsB protein.	168	8.30E-36
168153_2	W22173	S.thermophilus exopolysaccharide synthesis operon ...	141	2.20E-31
168153_2	W14074	S.thermophilus exopolysaccharide biosynthesis enzy...	141	2.20E-31
168153_2	P70458	Sequence of gpD encoded by segment of Xanthomonas ...	183	2.30E-30
168153_1	W22175	S.thermophilus exopolysaccharide synthesis operon ...	141	6.40E-35
168153_1	W14076	S.thermophilus exopolysaccharide biosynthesis enzy...	141	9.50E-35
168153_1	W22174	S.thermophilus exopolysaccharide synthesis operon ...	162	9.50E-30
168153_1	W14075	S.thermophilus exopolysaccharide biosynthesis enzy...	162	9.50E-30

168339_2	W27736	Putative O-antigen transporter protein.	820	5.70E-11.5
<b>GenBank</b>				
HGS001	gnlPIDle1183136	similar to 3-oxoacyl- acyl-carrier protein	569	2.20E-129
HGS001	gil151943	ORF3; putative [Rhodobacter capsulatus]	404	1.40E-92
HGS001	gil2983572	(AE000723) 3-oxoacyl-[acyl-carrier-protein	311	5.10E-92
HGS001	gil1276662	beta-ketoacyl-acyl carrier protein synthase	292	3.90E-90
HGS001	gil2313291	(AE000540) beta-ketoacyl-acyl carrier protein	269	3.50E-89
HGS001	gnlPIDle1183019	similar to 3-oxoacyl- acyl-carrier protein	373	2.00E-86
HGS001	gil1143069	3-ketoacyl carrier protein synthase III	287	3.60E-86
HGS001	gil22744	beta-ketoacyl-acyl carrier protein synthase	292	1.20E-85
HGS001	gil311686	3-ketoacyl-acyl carrier protein synthase	322	3.40E-85
HGS001	gil145898	beta-ketoacyl-acyl carrier protein synthase	366	7.30E-84
HGS002	gil142833	ORF2 [Bacillus subtilis] >gnlPIDle11851...	215	2.50E-70
HGS002	gnlPIDle1019368	hypothetical protein [Synechocystis sp.]	235	8.50E-67
HGS002	gil2983165	(AE000694) UDP-N-acetylglucosamine...	207	1.10E-58
HGS002	gil404010	ORF2 [Bacillus licheniformis] >pir14022...	251	1.10E-50
HGS002	gil2688520	(AE001161) UDP-N-acetylglucosamine dehydrog...	197	1.80E-42
HGS002	gil1841789	UDP-N-acetylglucosamine reductase...	249	7.10E-40
HGS002	gil2983149	(AE000693) UDP-N-acetylglucosamine...	212	3.80E-36
HGS002	gil431730	UDP-N-acetylglucosamine reductase...	119	4.50E-22
HGS002	gil1573234	UDP-N-acetylglucosamine reductase...	139	6.20E-22
HGS002	gil290456	UDP-N-acetylglucosamine reductase...	123	2.90E-20
HGS003	gnlPIDle1183192	similar to enoyl- acyl-carrier protein r...	743	1.80E-97
HGS003	gil142010	Shows 70.2% similarity and 48.6% identit...	519	8.90E-80
HGS003	gnlPIDle1017769	enoyl-[acyl-carrier-protein] reductase [...]	482	2.10E-73
HGS003	gil2313282	(AE000539) enoyl-(acyl-carrier-protein) ...	449	1.70E-71
HGS003	gil145851	envM [Escherichia coli] >gil587106 enoyl...	388	3.70E-71
HGS003	gil153955	envM protein [Salmonella typhimurium] >p...	386	2.10E-69
HGS003	gil1574591	short chain alcohol dehydrogenase homolo...	362	3.10E-68
HGS003	gil2983915	(AE000745) enoyl-[acyl-carrier-protein] ...	268	1.10E-64
HGS003	gil1053075	orf1; similar to E.coli EnvM [Proteus mi...	259	2.60E-29
HGS003	gnlPIDle1188732	(AJ003124) enoyl-ACP reductase [Petunia ...]	154	2.20E-28
HGS004	gnlPIDle276830	UDP-N-acetylglucosamine 1-carboxyvinyltr...	1251	2.50E-195

HGS004	gil1415662	UDP-N-acetylglucosamine 1-carboxyvinyl t...	534	1.40E-139
HGS004	gnllPID1d1010850	UDP-N-acetylglucosamine 1-carboxyvinyltr...	732	7.50E-138
HGS004	gil141344	UDP-N-acetylglucosamine 1-carboxyvinyltr...	537	2.90E-137
HGS004	gil1574635	UDP-N-acetylglucosamine enolpyruvyl tran...	536	4.70E-136
HGS004	gil146902	UDP-N-acetylglucosamine enolpyruvyl tran...	509	5.10E-134
HGS004	gil2983705	(AE000732) UDP-N-acetylglucosamine 1-car...	492	6.20E-121
HGS004	gnllPID1e229797	UDP-N-acetylglucosamine enolpyruvyl tran...	606	3.00E-119
HGS004	gil699337	UDP-N-acetylglucosamine 1-carboxyvinyl tr...	605	1.10E-118
HGS004	gil2313767	(AE000578) UDP-N-acetylglucosamine enolp...	440	1.90E-117
HGS005	gil143434	Rho Factor [Bacillus subtilis]	755	1.10E-190
HGS005	gil853769	transcriptional terminator Rho [Bacillus ...	746	1.80E-189
HGS005	gil2983405	(AE000711) transcriptional terminator Rho...	580	2.10E-154
HGS005	gil454859	The first ATG in the open reading frame ...	543	7.90E-150
HGS005	gil147607	transcription termination factor [Escheri...	592	9.40E-149
HGS005	gil49363	ho Factor [Salmonella typhimurium] >pirl...	592	1.70E-148
HGS005	gnllPID1e220353	Rho gene product [Streptomyces lividans] ...	575	4.90E-148
HGS005	gil1573263	transcription termination factor rho (rho...	575	5.40E-147
HGS005	gil49365	Rho factor [Neisseria gonorrhoeae] >pirl...	590	1.40E-146
HGS005	gil2313666	(AE000569) transcription termination fact...	547	8.10E-146
HGS006	gil580904	homologous to E.coli rnpA [Bacillus subt...	295	8.10E-37
HGS006	gnllPID1d1005777	protein component of ribonuclease P [Bac...	293	1.60E-36
HGS006	gnllPID1d1004132	RNaseP C5 subunit [Mycoplasma capricolum...	99	3.60E-22
HGS006	gil144147	rnpA [Buchnera aphidicola] >gil2827012 (...	97	3.90E-10
HGS006	gil511457	RNase P protein component [Coxiella burn...	117	2.30E-09
HGS007M	gnllPID1d1005718	replicative DNA helicase [Bacillus subti...	579	6.20E-169
HGS007M	gil3282821	(AF045058) DnaC replicative helicase [Ba...	536	3.60E-156
HGS007M	gnllPID1e321938	helicase [Rhodothermus marinus]	433	1.50E-123
HGS007M	gil2335167	(AF006675) DNA helicase [Rhodothermus ma...	271	2.90E-109
HGS007M	gnllPID1e211889	DNA-replication helicase [Odontella sine...	395	1.60E-108
HGS007M	gnllPID1e1263993	(AL022118) replicative DNA helicase DnaB...	235	3.20E-103
HGS007M	gnllPID1e244747	gene 40 [Bacteriophage SPP1] >gil529650 ...	477	4.40E-103
HGS007M	gil2983861	(AE000742) replicative DNA helicase [Aqua...	244	1.10E-102
HGS007M	gil2314528	(AE000636) replicative DNA helicase (dna...	246	7.70E-101

HGS007M	gnlIPIDd1011167	replicative DNA helicase [Synechocystis ...	209	1.50E-100
HGS008	gnlIPIDe1185181	malonyl CoA-acyl carrier protein transac...	560	4.30E-90
HGS008	gil1502420	malonyl-CoA:Acyl carrier protein transac...	391	1.40E-86
HGS008	gil3282803	(AF044668) malonyl CoA-acyl carrier prot...	308	2.50E-75
HGS008	gil2738154	malonyl-CoA:acyl carrier protein transac...	283	3.40E-75
HGS008	gil145887	malonyl coenzyme A-acyl carrier protein ...	304	6.30E-75
HGS008	gil1573113	malonyl coenzyme A-acyl carrier protein ...	270	7.60E-74
HGS008	gil2983416	(AE000712) malonyl-CoA:Acyl carrier prot...	213	2.70E-73
HGS008	gil840626	transacylase [Bacillus subtilis]	221	1.20E-66
HGS008	gil3150402	(AC004165) putative malonyl-CoA:Acyl car...	235	1.60E-57
HGS008	gnlIPIDe1185300	pksC [Bacillus subtilis] >gnlIPIDe11833...	145	4.40E-38
HGS009	gil460911	fructose-bisphosphate aldolase [Bacillus...	1169	2.10E-154
HGS009	gnlIPIDe1251871	fructose-1,6-bisphosphate aldolase type ...	1121	6.70E-148
HGS009	gnlIPIDd1003809	hypothetical protein [Bacillus subtilis]...	467	1.50E-110
HGS009	gil2313265	(AE000538) fructose-bisphosphate aldolas...	252	6.40E-91
HGS009	gil1673788	(AE000015) Mycoplasma pneumoniae, fructo...	238	4.60E-81
HGS009	gil1045692	fructose-bisphosphate aldolase [Mycoplas...	226	6.40E-77
HGS009	gnlIPIDd1016691	Tagatose-bisphosphate aldolase GatY (EC ...	279	2.30E-75
HGS009	gil599738	unknown function [Escherichia coli] >pir...	274	2.00E-74
HGS009	gil1732204	putative aldolase [Vibrio furnissii]	277	5.00E-74
HGS009	gil606077	ORF_o286 [Escherichia coli] >gil1789526 ...	264	1.30E-73
HGS014	gil40100	rodC (tag3) polypeptide (AA 1-746) [Baci...	597	1.70E-86
HGS014	gnlIPIDe1169895	tasA [Streptococcus pneumoniae]	108	4.90E-27
HGS014	gil2621425	(AE000822) teichoic acid biosynthesis pr...	142	2.00E-23
HGS014	gil2621421	(AE000822) teichoic acid biosynthesis pr...	147	5.90E-22
HGS014	gil143725	putative [Bacillus subtilis] >gnlIPIDe1...	114	4.60E-19
HGS014	gil547513	orf3 [Haemophilus influenzae] >pir1S4924...	106	5.60E-14
HGS014	gnlIPIDd1027517	(AB009477) 395aa long hypothetical prote...	79	4.20E-12
HGS014	gil2072447	EpsJ [Lactococcus lactis cremoris]	106	5.20E-10
HGS014	gil915199	ggaB [Bacillus subtilis] >gnlIPIDe11844...	89	8.10E-08
HGS016	gnlIPIDe276830	UDP-N-acetylglucosamine 1-carboxyvinyltr...	1251	2.50E-195
HGS016	gil415662	UDP-N-acetylglucosamine 1-carboxyvinyl t...	534	1.40E-139
HGS016	gnlIPIDd1010850	UDP-N-acetylglucosamine 1-carboxyvinyltr...	732	7.50E-138

HGS016	gil141344	UDP-N-acetylglucosamine 1-carboxyvinyltr...	537	2.90E-137
HGS016	gil1574635	UDP-N-acetylglucosamine enolpyruvyl tran...	536	4.70E-136
HGS016	gil146902	UDP-N-acetylglucosamine enolpyruvyl tran...	509	5.10E-134
HGS016	gil2983705	(AE000732) UDP-N-acetylglucosamine 1-car...	492	6.20E-121
HGS016	gnlPIDle229797	UDP-N-acetylglucosamine enolpyruvyl tran...	606	3.00E-119
HGS016	gil699337	UDP-N-acetylglucosamine 1-carboxyvinyl tr...	605	1.10E-118
HGS016	gil2313767	(AE000578) UDP-N-acetylglucosamine enolp...	440	1.90E-117
HGS018	gnlPIDle1182642	similar to DNA ligase [Bacillus subtilis...	1574	9.60E-287
HGS018	gnlPIDle1017321	DNA ligase [Synechocystis sp.] >pirS744...	830	5.70E-209
HGS018	gil1574651	DNA ligase (lig) [Haemophilus influenzae...	484	1.30E-204
HGS018	gil607820	DNA ligase [Rhodothermus marinus] >spIP4...	833	1.60E-204
HGS018	gil155088	DNA ligase [Thermus aquaticus thermophil...	428	3.10E-201
HGS018	gil609276	DNA ligase [Thermus scotoductus] >pirS5...	436	1.10E-200
HGS018	gil2983242	(AE000699) DNA ligase (NAD dependent) [A...	724	1.00E-179
HGS018	gil49284	DNA ligase [Zymomonas mobilis] >pirS206...	523	1.60E-170
HGS018	gnlPIDle1237759	(AL021287) DNA ligase [Mycobacterium tub...	529	1.80E-161
HGS018	gnlPIDle349403	DNA ligase [Mycobacterium leprae]	527	7.30E-160
HGS019	dbjID86417_12	YfiG [Bacillus subtilis] >gnlPIDle11827...	559	8.00E-72
HGS019	gil1044986	methionine aminopeptidase [Bacillus subt...	254	4.50E-58
HGS019	gil1574578	methionine aminopeptidase (map) [Haemoph...	185	5.10E-56
HGS019	gnlPIDle1172953	(AL008883) methionine aminopeptidase [My...	214	1.10E-51
HGS019	gil2982825	(AE000672) methionyl aminopeptidase [Aqu...	192	3.70E-48
HGS019	gnlPIDle1253272	(AL021958) methionine aminopeptidase [My...	130	5.20E-48
HGS019	gil2687996	(AE001123) methionine aminopeptidase (ma...	195	9.00E-48
HGS019	gnlPIDle1254451	methionine aminopeptidase [Streptomyces ...	151	2.10E-43
HGS019	gil975723	methionine aminopeptidase I [Saccharomyc...	294	3.60E-43
HGS019	gil2583129	(AC002387) putative methionine aminopept...	211	2.10E-41
HGS022	gnlPIDle1182648	alternate gene name: yedB; similar to am...	1586	2.80E-212
HGS022	gil2589195	(AF008553) Glu-tRNA Gln amidotransferase ...	1436	1.70E-198
HGS022	gnlPIDle1018331	amidase [Synechocystis sp.] >pirS772641...	867	2.30E-178
HGS022	gil2982954	(AE000680) glutamyl-tRNA (Gln) amidotran...	1247	6.50E-176
HGS022	gil1224069	amidase [Moraxella catarrhalis] >spIQ490...	522	4.40E-158
HGS022	gil2648182	(AE000943) Glu-tRNA amidotransferase, su...	548	1.30E-145

HGS022	gnlIPIDle349405	probable amidase [Mycobacterium leprae]	465	6.30E-143
HGS022	gnlIPIDle1237756	(AL021287) putative Glu-tRNA-Gln amidotr...	470	1.90E-141
HGS022	gil2313964	(AE000594) amidase [Helicobacter pylori]...	550	7.30E-123
HGS022	gil2622613	(AE000910) amidase [Methanobacterium the...	524	5.80E-116
HGS023	gil1354211	PET112-like protein [Bacillus subtilis] ...	2291	2.90E-307
HGS023	gil2653657	Bacillus subtilis PET112-like protein [B...	1313	1.20E-250
HGS023	gil2589196	(AF008553) Glu-tRNA <sup>Gln</sup> amidotransferase ...	1315	4.20E-250
HGS023	gnlIPIDle1182649	similar to pet112-like protein [Bacillus...	1346	7.10E-224
HGS023	gil2983123	(AE000691) glutamyl-tRNA (Gln) amidotran...	931	2.30E-165
HGS023	gnlIPIDd1019042	PET112 [Synechocystis sp.] >pirIS75850IS...	859	4.10E-161
HGS023	gil1224071	unknown [Moraxella catarrhalis] >spIQ490...	323	3.90E-132
HGS023	gil2313783	(AE000579) PET112-like protein [Helicoba...	664	6.80E-132
HGS023	gil2688237	(AE001140) glu-tRNA amidotransferase, su...	318	4.00E-131
HGS023	gil1590917	Glu-tRNA amidotransferase (gatB) [Methan...	263	8.60E-125
HGS024	gil2465557	(AF011545) YedA [Bacillus subtilis] >gil...	237	6.30E-27
HGS024	gnlIPIDd1011444	hypothetical protein [Synechocystis sp.]...	153	8.60E-22
HGS024	gil2648183	(AE000943) Glu-tRNA amidotransferase, su...	126	1.80E-21
HGS024	gnlIPIDle1237757	(AL021287) putative Glu-tRNA-Gln amidotr...	166	1.80E-17
HGS024	gil2984354	(AE000775) glutamyl-tRNA (Gln) amidotran...	102	2.70E-17
HGS024	gnlIPIDle349616	hypothetical protein MLCB637.12 [Mycobac...	154	7.10E-16
HGS025	gnlIPIDd1005830	stage V sporulation [Bacillus subtilis] ...	496	4.90E-69
HGS025	gnlIPIDd1011124	peptidyl-tRNA hydrolase [Synechocystis s...	307	2.10E-49
HGS025	gil2983032	(AE000685) peptidyl-tRNA hydrolase [Aqui...	386	2.20E-49
HGS025	gnlIPIDle304565	Pth [Mycobacterium tuberculosis] >gnlPL...	266	2.60E-43
HGS025	gil1045760	peptidyl-tRNA hydrolase homolog [Mycopla...	211	1.40E-39
HGS025	gil2314676	(AE000648) peptidyl-tRNA hydrolase (pth)...	102	3.30E-39
HGS025	gil1674312	(AE000058) Mycoplasma pneumoniae, peptid...	208	9.50E-39
HGS025	gil1127571	peptidyl-tRNA hydrolase [Chlamydia trach...	187	7.00E-37
HGS025	gil1573366	peptidyl-tRNA hydrolase (pth) [Haemophil...	201	8.50E-34
HGS025	gil581202	peptidyl-tRNA hydrolase [Escherichia col...	186	2.50E-27
HGS026	gil853776	peptide chain release factor 1 [Bacillus...	889	6.10E-160
HGS026	gnlIPIDd1009421	Peptide Termination Factor [Mycoplasma c...	715	1.10E-126
HGS026	gnlIPIDd1019559	peptide chain release factor [Synechocys...	539	2.70E-121

HGS026	gil2688096	(AE001130) peptide chain release factor ...	627	1.80E-115
HGS026	gnlPID1d1015453	Peptide chain release factor 1 (RF-1) [E...	467	3.90E-113
HGS026	gil968930	peptide chain release factor 1 [Escheric...	463	1.30E-112
HGS026	gil147567	peptide chain release factor 1 [Escheric...	467	3.40E-112
HGS026	gil154104	release factor 1 [Salmonella typhimurium...	460	2.90E-111
HGS026	gil1574404	polypeptide chain release factor 1 (prfA...	449	1.50E-109
HGS026	gil2313158	(AE000529) peptide chain release factor ...	576	1.20E-104
HGS028	gil2331287	(AF013188) release factor 2 [Bacillus...	769	2.50E-173
HGS028	spiP28367IRF2_BACSU	PEPTIDE CHAIN RELEASE FACTOR 2 (RF-2)...	742	3.00E-157
HGS028	gil2984119	(AE000758) peptide chain release fact...	442	2.20E-128
HGS028	gnlPID1e254636	peptide release factor 2 [Bacillus fi...	718	2.90E-125
HGS028	pirIS76448IS76448	translation releasing factor RF-2 - S...	883	3.30E-116
HGS028	pirA64190IA64190	translation releasing factor RF-2 - H...	444	1.70E-110
HGS028	gil154276	peptide chain release factor 2 [Salmo...	444	1.80E-108
HGS028	gil2687953	(AE001120) peptide chain release fact...	408	3.90E-108
HGS028	gil2367172	(AE000372) peptide chain release fact...	437	1.60E-107
HGS028	gil147569	peptide chain release factor 2 [Esche...	434	4.00E-107
HGS030	gnlPID1d1005806	unknown [Bacillus subtilis] >gnlPID1e1...	283	2.60E-64
HGS030	gil1317687	(AF065312) thymidylate kinase [Yersinia ...	124	3.00E-43
HGS030	gil2983484	(AE000716) thymidylate kinase [Aquifex a...	272	2.40E-37
HGS030	gil1244710	thymidylate kinase [Escherichia coli] >g...	136	7.20E-34
HGS030	gil2650584	(AE001102) thymidylate kinase (tmk) [Arc...	71	2.60E-30
HGS030	gil1045674	thymidylate kinase [Mycoplasma genitaliu...	173	8.20E-28
HGS030	gil1673808	(AE000016) Mycoplasma pneumoniae, thymid...	171	1.70E-27
HGS030	gil1246364	thymidylate:zeocin resistance protein:ND...	136	2.20E-27
HGS030	gil1246361	thymidine:thymidylate kinase:zeocin resi...	136	4.30E-27
HGS030	gil950071	ATP-bind. pyrimidine kinase [Mycoplasma ...	80	8.70E-21
HGS031	gnlPID1e1185242	uridylate kinase [Bacillus subtilis] >pi...	920	8.40E-123
HGS031	gnlPID1d1019291	uridine monophosphate kinase [Synecocys...	530	1.70E-96
HGS031	gnlPID1e1296663	(AL023797) uridylate kinase [Streptomyce...	678	2.10E-89
HGS031	gnlPID1e248883	hypothetical protein MTCY274.14c [Mycoba...	416	6.00E-89
HGS031	gnlPID1e327783	uridylate kinase [Mycobacterium leprae]	403	7.90E-86
HGS031	gil473234	uridine 5'-monophosphate (UMP) kinase [E...	384	2.10E-72

HGS031	gil1552748	uridine 5'-monophosphate (UMP) kinase [E...	375	3.60E-71
HGS031	gil1574616	mukB suppressor protein (smbA) [Haemophi...	409	3.70E-71
HGS031	gil2983290	(AE000703) UMP kinase [Aquifex aeolicus]	452	3.70E-58
HGS031	gil1518662	UMP kinase [Chlamydia trachomatis] >spP...	323	9.10E-55
HGS032	gil755152	highly hydrophobic integral membrane pro...	297	2.40E-81
HGS032	gil1235660	RfbA [Myxococcus xanthus] >sp Q50862 IRFB...	173	4.90E-24
HGS032	gnllPIDd1017629	ABC transporter [Synechocystis sp.] >pir...	149	1.50E-19
HGS032	gnllPIDd1029275	(AB010294) integral membrane component o...	126	6.40E-19
HGS032	gnllPIDd1008332	putative integral membrane component of ...	125	9.10E-19
HGS032	gnllPIDd1029271	(AB010293) integral membrane component o...	125	9.10E-19
HGS032	gnllPIDd1029279	(AB010295) integral membrane component o...	125	9.10E-19
HGS032	gnllPIDd1029264	(AB010150) integral membrane component o...	109	3.00E-15
HGS032	gil2983575	(AE000723) ABC transporter (ABC-2 subfam...	71	9.60E-13
HGS032	gil609595	homologous to kpsM (E.coli), bexB (H.inf...	78	2.60E-12
HGS033	gil755153	ATP-binding protein [Bacillus subtilis] ...	655	9.30E-94
HGS033	gil609596	ATP-binding protein [Serratia marcescens]	387	3.70E-69
HGS033	gil765059	ABC-transporter protein [Klebsiella pneu...	371	3.70E-69
HGS033	gil567183	ATP-binding protein [Klebsiella pneumoni...	367	1.20E-67
HGS033	gil304013	abcA [Aeromonas salmonicida] >pir A36918...	294	7.20E-59
HGS033	gnllPIDd1020415	(AB002668) ABC transport protein [Actino...	323	4.00E-57
HGS033	gil1123030	CpxA [Actinobacillus pleuropneumoniae]	190	2.40E-56
HGS033	gil3135679	(AF064070) putative ABC-2 transporter hy...	219	2.10E-53
HGS033	gil2983576	(AE000723) ABC transporter [Aquifex aeol...	294	2.10E-53
HGS033	gil1235661	RfbB [Myxococcus xanthus] >sp Q50863 IRFB...	336	6.70E-53
HGS034	gil143467	ribosomal protein S4 [Bacillus subtilis]...	798	4.50E-106
HGS034	gil2314460	(AE000633) ribosomal protein S4 (rps4) [...	322	1.50E-62
HGS034	gil2982819	(AE000672) ribosomal protein S04 [Aquife...	253	2.00E-62
HGS034	gil606231	30S ribosomal subunit protein S4 [Escher...	292	2.40E-58
HGS034	gnllPIDe1234848	(AJ223236) ribosomal protein S4 [Salmone...	292	6.10E-58
HGS034	gil1573812	ribosomal protein S4 (rpS4) [Haemophilus...	292	1.60E-57
HGS034	gil639791	ribosomal protein S4 [Mycoplasma pneumon...	260	1.90E-56
HGS034	gil1046011	ribosomal protein S4 [Mycoplasma genital...	245	2.10E-54
HGS034	gnllPIDe1316061	RpsD [Mycobacterium tuberculosis] >gnllP...	270	1.40E-52



HGS034	gii144143	ribosomal protein S4 [Buchnera aphidicol...	255	2.00E-51
HGS036	gi2648781	(AE000980) dipeptide ABC transporter, AT...	136	1.90E-40
HGS036	gnlPIDle1264523	(AL022121) putative peptide ABC transpor...	185	5.50E-35
HGS036	gii143607	sporulation protein [Bacillus subtilis]	191	7.70E-34
HGS036	gnlPIDle1183166	oligopeptide ABC transporter (ATP-bindin...	191	7.70E-34
HGS036	gnlPIDle1253461	oligopeptide transport ATP-binding prote...	213	5.50E-33
HGS036	gi2313342	(AE000544) oligopeptide ABC transporter,...	258	7.60E-32
HGS036	gnlPIDld1015858	Dipeptide transport ATP-binding protein ...	205	1.10E-31
HGS036	gi47346	AmiE protein [Streptococcus pneumoniae] ...	202	7.40E-31
HGS036	gi972897	DppD [Haemophilus influenzae] >gil157411...	204	1.40E-30
HGS036	gi677943	AppD [Bacillus subtilis] >gnlPIDle11831...	205	9.70E-30
HGS040	gnlPIDle1185713	elongation factor P [Bacillus subtilis] ...	702	7.00E-91
HGS040	gi1399829	elongation factor P [Synecococcus PCC79...	541	4.90E-69
HGS040	gnlPIDld1010902	elongation factor P [Synecocystis sp.] ...	535	3.20E-68
HGS040	gi951349	ORF1; putative [Anabaena sp.] >sp Q44247...	505	3.80E-64
HGS040	gnlPIDle290977	unknown [Mycobacterium tuberculosis] >gn...	480	9.20E-61
HGS040	gnlPIDle1169516	elongation factor P [Corynebacterium glu...	460	4.80E-58
HGS040	gi2983772	(AE000736) elongation factor P [Aquifex ...	435	1.10E-54
HGS040	gi1658506	elongation factor P homologue; EF-P [Bac...	203	7.20E-52
HGS040	gi2313266	(AE000538) translation elongation factor...	409	4.00E-51
HGS040	gi536991	elongation factor P [Escherichia coli] >...	362	9.40E-45
168153_3	gnlPIDld1028815	(AB009524) Vi polysaccharide biosynthes...	237	5.80E-72
168153_3	gi47961	wcdB; ORF3 in citation [1] [Salmonella ...	234	1.80E-71
168153_3	gi1590951	UDP-glucose 4-epimerase (galE) [Methano...	148	3.20E-60
168153_3	pirC69149IC69149	conserved hypothetical protein MTH380 ~...	151	1.90E-50
168153_3	gi1143204	ORF2; Method: conceptual translation s...	227	4.50E-47
168153_3	gnlPIDle316552	unknown [Mycobacterium tuberculosis] >g...	109	4.70E-45
168153_3	gnlPIDle1185960	similar to NDP-sugar epimerase [Bacillu...	155	1.80E-39
168153_3	gnlPIDle1289548	(AL023093) putative sugar dehydratase [M...	86	1.80E-36
168153_3	gnlPIDle288124	glucose epimerase [Bacillus thuringiensis]	95	2.70E-35
168153_3	gi1591707	capsular polysaccharide biosynthesis pr...	85	1.60E-34
168153_2	gnlPIDle1184467	alternate gene name: yvhA [Bacillus subt...	354	4.90E-45
168153_2	gi1657652	Cap8M [Staphylococcus aureus]	138	9.00E-42
168153_2	gi11773352	Cap5M [Staphylococcus aureus]	138	9.00E-42

l68153_2	gnllPIDle238668	hypothetical protein [Bacillus subtilis]...	139	6.10E-39
l68153_2	gil1199573	spsB [Shingomonas sp.] >gil1314578 gluc...	168	4.40E-35
l68153_2	gnllPIDld1005318	ORF14 [Klebsiella pneumoniae] >spQ48460...	260	5.50E-33
l68153_2	gnllPIDld1020425	(AB002668) galactosyltransferase [Actino...	155	5.60E-33
l68153_2	gnllPIDld1029082	(AB010415) glycosyltransferase [Actinoba...	155	2.00E-32
l68153_2	gnllPIDld1019174	galactosyl-1-phosphate transferase [Syn...	139	2.30E-32
l68153_2	gnllPIDle220381	structural gene [Agrobacterium radiobacter]	138	2.40E-32
l68153_1	gil1276880	EpsG [Streptococcus thermophilus]	141	3.40E-34
l68153_1	gil1276879	EpsF [Streptococcus thermophilus]	162	1.70E-29
l68153_1	gil633699	WbcQ [Yersinia enterocolitica] >pirIS512...	134	9.10E-26
l68153_1	gnllPIDle238704	hypothetical protein [Bacillus subtilis]...	131	1.90E-18
l68153_1	gil2983976	(AE000749) capsular polysaccharide biosy...	134	1.50E-15
l68153_1	gnllPIDld1005311	ORF7 [Klebsiella pneumoniae] >spQ48453l...	94	2.10E-12
l68153_1	gil633696	WbcN [Yersinia enterocolitica] >pirIS512...	123	2.50E-12
l68153_1	gil755606	unknown [Bacillus subtilis]	144	5.40E-12
l68153_1	gil1146237	21.4% of identity to trans-acting transc...	144	6.00E-12
l68153_1	gnllPIDle238664	hypothetical protein [Bacillus subtilis]...	141	3.20E-11
l68339_2	gnllPIDle1169894	putative repeating unit transporter ...	234	5.70E-57
l68339_2	gil2209215	(AF004325) putative oligosaccharide ...	139	4.90E-37
l68339_2	gil633692	Wzx [Yersinia enterocolitica] >pirIS...	141	3.00E-31
l68339_2	gil2621404	(AE000819) O-antigen transporter [Me...	129	8.90E-29
l68339_2	gil2072448	EpsK [Lactococcus lactis cremoris]	199	4.00E-27
l68339_2	spIP37746IRFBX_ECOLI	PUTATIVE O-ANTIGEN TRANSPORTER.	140	2.10E-23
l68339_2	gnllPIDld1016603	Putative O-antigen transporter. [Esc...	140	2.90E-23
l68339_2	gil510252	membrane protein [Escherichia coli]	140	8.10E-23
l68339_2	gil2621427	(AE000822) O-antigen transporter [Me...	122	3.10E-20
l68339_2	gil152778	RFBX [Shigella dysenteriae] >pirIS34...	114	8.50E-19

**[0090]** If the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends, relative to the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This corrected score is what is used for the purposes of the present invention. Only nucleotides outside the 5' and 3' nucleotides of the subject sequence, as displayed by the FASTDB alignment, which are not matched/aligned with the query sequence, are calculated for the purposes of manually adjusting the percent identity score.

**[0091]** For example, a 90 nucleotide subject sequence is aligned to a 100 nucleotide query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignment of the first 10 nucleotides at 5' end. The 10 unpaired nucleotides represent 10% of the sequence (number of nucleotides at the 5' and 3' ends not matched/total number of nucleotides in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 nucleotides were perfectly matched the final percent identity would be 90%. In another example, a 90 nucleotide subject sequence is compared with a 100 nucleotide query sequence. This time the deletions are internal deletions so that there are no nucleotides on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only nucleotides 5' and 3' of the subject sequence which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are made for the purposes of the present invention.

**[0092]**     *Vectors and Host Cell*

**[0093]**     The present invention also relates to vectors which include the isolated DNA molecules of the present invention, host cells comprising the recombinant vectors, and the production of *S. aureus* polypeptides and peptides of the present invention expressed by the host cells.

**[0094]**     Recombinant constructs may be introduced into host cells using well known techniques such as infection, transduction, transfection, transvection, electroporation and transformation. The vector may be, for example, a phage, plasmid, viral or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

**[0095]**     The *S. aureus* polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged *in vitro* using an appropriate packaging cell line and then transduced into host cells.

**[0096]**     Preferred are vectors comprising *cis*-acting control regions to the polynucleotide of interest. Appropriate *trans*-acting factors may be supplied by the host, supplied by a complementing vector or supplied by the vector itself upon introduction into the host.

**[0097]**     In certain preferred embodiments in this regard, the vectors provide for specific expression, which may be inducible and/or cell type-specific. Particularly preferred among such vectors are those inducible by environmental factors that are easy to manipulate, such as temperature and nutrient additives.

**[0098]**     Expression vectors useful in the present invention include chromosomal-, episomal- and virus-derived vectors, *e.g.*, vectors derived from bacterial plasmids, bacteriophage, yeast episomes, yeast chromosomal elements, viruses such as baculoviruses, papova viruses, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as cosmids and phagemids.

**[0099]**     The DNA insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the *E. coli lac*, *trp*, *phoA* and *tac* promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination and, in the transcribed region, a ribosome binding site for translation. The coding portion of the mature transcripts expressed by the constructs will

preferably include a translation initiating site at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

**[0100]** As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin, or ampicillin resistance genes for culturing in *E. coli* and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as *E. coli*, *Streptomyces* and *Salmonella typhimurium* cells; fungal cells, such as yeast cells (e.g., *Saccharomyces cerevisiae* or *Pichia pastoris* (ATCC Accession No. 201178)); insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

**[0101]** Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE9, pQE10 available from Qiagen; pBS vectors, Phagescript vectors, Bluescript vectors, pNH8A, pNH16a, pNH18A, pNH46A available from Stratagene Cloning Systems, Inc.; pET series of vectors available from Novagen; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech, Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Preferred expression vectors for use in yeast systems include, but are not limited to pYES2, pYD1, pTEF1/Zeo, pYES2/GS, pPICZ, pGAPZ, pGAPZalph, pPIC9, pPIC3.5, pHIL-D2, pHIL-S1, pPIC3.5K, pPIC9K, and PAO815 (all available from Invitrogen, Carlsbad, CA). Other suitable vectors will be readily apparent to the skilled artisan.

**[0102]** Among known bacterial promoters suitable for use in the present invention include the *E. coli* *lacI* and *lacZ* promoters, the T3, T5 and T7 promoters, the *gpt* promoter, the lambda PR and PL promoters and the *trp* promoter. Suitable eukaryotic promoters include the CMV immediate early promoter, the HSV thymidine kinase promoter, the early and late SV40 promoters, the promoters of retroviral LTRs, such as those of the Rous sarcoma virus (RSV), and metallothionein promoters, such as the mouse metallothionein-I promoter.

**[0103]** Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection or other methods. Such methods are described in many standard laboratory manuals (for example, Davis, *et al.*, *Basic Methods In Molecular Biology* (1986)). It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

**[0104]** Transcription of DNA encoding the polypeptides of the present invention by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are *cis*-acting elements of DNA, usually about from 10 to 300 nucleotides that act to increase transcriptional activity of a promoter in a given host cell-type. Examples of enhancers include the SV40 enhancer, which is located on the late side of the replication origin at nucleotides 100 to 270, the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

**[0105]** For secretion of the translated polypeptide into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the expressed polypeptide, for example, the amino acid sequence KDEL. The signals may be endogenous to the polypeptide or they may be heterologous signals.

**[0106]** The polypeptide may be expressed in a modified form, such as a fusion protein, and may include not only secretion signals, but also additional heterologous functional regions. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence in the host cell, during purification, or during subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to polypeptides to engender secretion or excretion, to improve stability and to facilitate purification, among others, are familiar and routine techniques in the art. A preferred fusion protein comprises a heterologous region from immunoglobulin that is useful to solubilize proteins. For example, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is thoroughly advantageous for use in therapy and diagnosis and thus results, for example, in improved pharmacokinetic properties (EP-A 0232 262). On the other hand, for some uses it would be desirable to be able to delete the Fc part after the fusion protein has been expressed, detected and purified in the advantageous manner described. This is the case when Fc portion proves to be a hindrance to use in therapy and diagnosis, for example when the fusion protein is to be used as antigen for immunizations. In drug discovery, for example, human proteins, such as, hIL5-receptor has been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. See Bennett, D. et al. (1995) J. Molec. Recogn. 8:52-58 and Johanson, K.

et al. (1995) J. Biol. Chem. 270 (16):9459-9471.

[0107] The *S. aureus* polypeptides can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography, lectin chromatography and high performance liquid chromatography ("HPLC") is employed for purification. Polypeptides of the present invention include naturally purified products, products of chemical synthetic procedures, and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect and mammalian cells.

[0108] Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

[0109] In one embodiment, the yeast *Pichia pastoris* is used to express any plasma membrane associated protein of the invention in a eukaryotic system. *Pichia pastoris* is a methylotrophic yeast which can metabolize methanol as its sole carbon source. A main step in the methanol metabolization pathway is the oxidation of methanol to formaldehyde using O<sub>2</sub>. This reaction is catalyzed by the enzyme alcohol oxidase. In order to metabolize methanol as its sole carbon source, *Pichia pastoris* must generate high levels of alcohol oxidase due, in part, to the relatively low affinity of alcohol oxidase for O<sub>2</sub>. Consequently, in a growth medium depending on methanol as a main carbon source, the promoter region of one of the two alcohol oxidase genes (*AOX1*) is highly active. In the presence of methanol, alcohol oxidase produced from the *AOX1* gene comprises up to approximately 30% of the total soluble protein in *Pichia pastoris*. See, Ellis, S.B., et al., *Mol. Cell. Biol.* 5:1111-21 (1985); Koutz, P.J., et al., *Yeast* 5:167-77 (1989); Tschopp, J.F., et al., *Nucl. Acids Res.* 15:3859-76 (1987). Thus, a heterologous coding sequence, such as, for example, a plasma membrane associated

polynucleotide of the present invention, under the transcriptional regulation of all or part of the *AOX1* regulatory sequence is expressed at exceptionally high levels in *Pichia* yeast grown in the presence of methanol.

[0110] In one example, the plasmid vector pPIC9K is used to express DNA encoding a plasma membrane associated polypeptide of the invention, as set forth herein, in a *Pichea* yeast system essentially as described in "*Pichia* Protocols: Methods in Molecular Biology," D.R. Higgins and J. Cregg, eds. The Humana Press, Totowa, NJ, 1998. This expression vector allows expression and secretion of a plasma membrane associated protein of the invention by virtue of the strong *AOX1* promoter linked to the *Pichia pastoris* alkaline phosphatase (PHO) secretory signal peptide (i.e., leader) located upstream of a multiple cloning site.

[0111] Many other yeast vectors could be used in place of pPIC9K, such as, pYES2, pYD1, pTEF1/Zeo, pYES2/GS, pPICZ, pGAPZ, pGAPZalpha, pPIC9, pPIC3.5, pHIL-D2, pHIL-S1, pPIC3.5K, and PAO815, as one skilled in the art would readily appreciate, as long as the proposed expression construct provides appropriately located signals for transcription, translation, secretion (if desired), and the like, including an in-frame AUG as required.

[0112] In another embodiment, high-level expression of a heterologous coding sequence, such as, for example, a plasma membrane associated polynucleotide of the present invention, may be achieved by cloning the heterologous polynucleotide of the invention into an expression vector such as, for example, pGAPZ or pGAPZalpha, and growing the yeast culture in the absence of methanol.

[0113] In addition to encompassing host cells containing the vector constructs discussed herein, the invention also encompasses host cells that have been engineered to delete or replace endogenous genetic material (e.g. coding sequences for the polypeptides of the present invention), and/or to include genetic material (e.g. heterologous polynucleotide sequences) that is operably associated with polynucleotides of the present invention, and which activates, alters, and/or amplifies endogenous polynucleotides of the present invention. For example, techniques known in the art may be used to operably associate heterologous control regions (e.g. promoter and/or enhancer) and endogenous polynucleotide sequences via homologous recombination (see, e.g. U.S. Patent No. 5,641,670, issued June 24, 1997; International Publication No. WO 96/29411, published September 26, 1996; International Publication No. WO 94/12650, published August 4, 1994; Koller et al., Proc. Natl. Acad. Sci. USA 86:8932-8935 (1989); and Zijlstra, et al., Nature 342:435-438 (1989), the disclosures of each of which



are hereby incorporated by reference in their entireties).

**[0114]** In addition, polypeptides of the invention can be chemically synthesized using techniques known in the art (e.g., see Creighton, 1983, *Proteins: Structures and Molecular Principles*, W.H. Freeman & Co., N.Y., and Hunkapiller et al., *Nature*, 310:105-111 (1984)). For example, a polypeptide corresponding to a fragment of a polypeptide can be synthesized by use of a peptide synthesizer. Furthermore, if desired, nonclassical amino acids or chemical amino acid analogs can be introduced as a substitution or addition into the polypeptide sequence. Non-classical amino acids include, but are not limited to, to the D-isomers of the common amino acids, 2,4-diáminobutyric acid, a-amino isobutyric acid, 4-aminobutyric acid, Abu, 2-amino butyric acid, g-Abu, e-Ahx, 6-amino hexanoic acid, Aib, 2-amino isobutyric acid, 3-amino propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, homocitrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine, b-alanine, fluoro-amino acids, designer amino acids such as b-methyl amino acids, Ca-methyl amino acids, Na-methyl amino acids, and amino acid analogs in general. Furthermore, the amino acid can be D (dextrorotary) or L (levorotary).

**[0115]** Non-naturally occurring variants may be produced using art-known mutagenesis techniques, which include, but are not limited to oligonucleotide mediated mutagenesis, alanine scanning, PCR mutagenesis, site directed mutagenesis (*see, e.g., Carter et al., Nucl. Acids Res.* 13:4331 (1986); and Zoller et al., *Nucl. Acids Res.* 10:6487 (1982)), cassette mutagenesis (*see, e.g., Wells et al., Gene* 34:315 (1985)), restriction selection mutagenesis (*see, e.g., Wells et al., Philos. Trans. R. Soc. London SerA* 317:415 (1986)).

**[0116]** The invention additionally, encompasses polypeptides of the present invention which are differentially modified during or after translation, such as for example, by glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand, etc. Any of numerous chemical modifications may be carried out by known techniques, including but not limited to: specific chemical cleavage by cyanogen bromide; trypsin; chymotrypsin; papain; V8 protease; NaBH<sub>4</sub>; acetylation; formylation; oxidation; reduction; and metabolic synthesis in the presence of tunicamycin, etc.

**[0117]** Additional post-translational modifications encompassed by the invention include, for example, N-linked or O-linked carbohydrate chains, processing of N-terminal or C-terminal ends, attachment of chemical moieties to the amino acid backbone, chemical modifications of N-linked or O-linked carbohydrate chains, and addition or deletion of an N-terminal

methionine residue as a result of procaryotic host cell expression. The polypeptides may also be modified with a detectable label, such as an enzymatic, fluorescent, isotopic or affinity label to allow for detection and isolation of the protein.

**[0118]** Also provided by the invention are chemically modified derivatives of the polypeptides of the invention which may provide additional advantages such as increased solubility, stability and circulating time of the polypeptide, or decreased immunogenicity (see U.S. Patent No. 4,179,337). The chemical moieties for derivitization may be selected from water soluble polymers such as polyethylene glycol, ethylene glycol/propylene glycol copolymers, carboxymethylcellulose, dextran, polyvinyl alcohol and the like. The polypeptides may be modified at random positions within the molecule, or at predetermined positions within the molecule and may include one, two, three or more attached chemical moieties.

**[0119]** The polymer may be of any molecular weight, and may be branched or unbranched. For polyethylene glycol, the preferred molecular weight is between about 1 kDa and about 100 kDa (the term "about" indicating that in preparations of polyethylene glycol, some molecules will weigh more, some less, than the stated molecular weight) for ease in handling and manufacturing. Other sizes may be used, depending on the desired therapeutic profile, which can include, for example, the duration of sustained release desired; the effects, if any, on biological activity; the ease in handling; the degree or lack of antigenicity; and other known effects of the polyethylene glycol on a therapeutic protein or analog. For example, the polyethylene glycol may have an average molecular weight of about 200, 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000, 6500, 7000, 7500, 8000, 8500, 9000, 9500, 10,000, 10,500, 11,000, 11,500, 12,000, 12,500, 13,000, 13,500, 14,000, 14,500, 15,000, 15,500, 16,000, 16,500, 17,000, 17,500, 18,000, 18,500, 19,000, 19,500, 20,000, 25,000, 30,000, 35,000, 40,000, 50,000, 55,000, 60,000, 65,000, 70,000, 75,000, 80,000, 85,000, 90,000, 95,000, or 100,000 kDa.

**[0120]** As noted above, the polyethylene glycol may have a branched structure. Branched polyethylene glycols are described, for example, in U.S. Patent No. 5,643,575; Morpurgo *et al.*, *Appl. Biochem. Biotechnol.* 56:59-72 (1996); Vorobjev *et al.*, *Nucleosides Nucleotides* 18:2745-2750 (1999); and Caliceti *et al.*, *Bioconjug. Chem.* 10:638-646 (1999), the disclosures of each of which are incorporated herein by reference in their entireties.

**[0121]** The polyethylene glycol molecules (or other chemical moieties) should be attached to the protein with consideration of effects on functional or antigenic domains of the protein.

There are a number of attachment methods available to those skilled in the art, e.g., EP 0 401 384, herein incorporated by reference (coupling PEG to G-CSF), see also Malik et al., *Exp. Hematol.* 20:1028-1035 (1992) (reporting pegylation of GM-CSF using tresyl chloride). For example, polyethylene glycol may be covalently bound through amino acid residues via a reactive group, such as, a free amino or carboxyl group. Reactive groups are those to which an activated polyethylene glycol molecule may be bound. The amino acid residues having a free amino group may include lysine residues and the N-terminal amino acid residues; those having a free carboxyl group may include aspartic acid residues glutamic acid residues and the C-terminal amino acid residue. Sulfhydryl groups may also be used as a reactive group for attaching the polyethylene glycol molecules. Preferred for therapeutic purposes is attachment at an amino group, such as attachment at the N-terminus or lysine group.

**[0122]** As suggested above, polyethylene glycol may be attached to proteins via linkage to any of a number of amino acid residues. For example, polyethylene glycol can be linked to a proteins via covalent bonds to lysine, histidine, aspartic acid, glutamic acid, or cysteine residues. One or more reaction chemistries may be employed to attach polyethylene glycol to specific amino acid residues (e.g., lysine, histidine, aspartic acid, glutamic acid, or cysteine) of the protein or to more than one type of amino acid residue (e.g., lysine, histidine, aspartic acid, glutamic acid, cysteine and combinations thereof) of the protein.

**[0123]** One may specifically desire proteins chemically modified at the N-terminus. Using polyethylene glycol as an illustration of the present composition, one may select from a variety of polyethylene glycol molecules (by molecular weight, branching, etc.), the proportion of polyethylene glycol molecules to protein (polypeptide) molecules in the reaction mix, the type of pegylation reaction to be performed, and the method of obtaining the selected N-terminally pegylated protein. The method of obtaining the N-terminally pegylated preparation (i.e., separating this moiety from other monopegylated moieties if necessary) may be by purification of the N-terminally pegylated material from a population of pegylated protein molecules. Selective proteins chemically modified at the N-terminus modification may be accomplished by reductive alkylation which exploits differential reactivity of different types of primary amino groups (lysine versus the N-terminal) available for derivatization in a particular protein. Under the appropriate reaction conditions, substantially selective derivatization of the protein at the N-terminus with a carbonyl group containing polymer is achieved.

**[0124]** As indicated above, pegylation of the proteins of the invention may be accomplished by any number of means. For example, polyethylene glycol may be attached to

the protein either directly or by an intervening linker. Linkerless systems for attaching polyethylene glycol to proteins are described in Delgado *et al.*, *Crit. Rev. Thera. Drug Carrier Sys.* 9:249-304 (1992); Francis *et al.*, *Intern. J. of Hematol.* 68:1-18 (1998); U.S. Patent No. 4,002,531; U.S. Patent No. 5,349,052; WO 95/06058; and WO 98/32466, the disclosures of each of which are incorporated herein by reference.

**[0125]** One system for attaching polyethylene glycol directly to amino acid residues of proteins without an intervening linker employs tresylated MPEG, which is produced by the modification of monmethoxy polyethylene glycol (MPEG) using tresylchloride ( $\text{ClSO}_2\text{CH}_2\text{CF}_3$ ). Upon reaction of protein with tresylated MPEG, polyethylene glycol is directly attached to amine groups of the protein. Thus, the invention includes protein-polyethylene glycol conjugates produced by reacting proteins of the invention with a polyethylene glycol molecule having a 2,2,2-trifluoroethane sulphonyl group.

**[0126]** Polyethylene glycol can also be attached to proteins using a number of different intervening linkers. For example, U.S. Patent No. 5,612,460, the entire disclosure of which is incorporated herein by reference, discloses urethane linkers for connecting polyethylene glycol to proteins. Protein-polyethylene glycol conjugates wherein the polyethylene glycol is attached to the protein by a linker can also be produced by reaction of proteins with compounds such as MPEG-succinimidylsuccinate, MPEG activated with 1,1'-carbonyldiimidazole, MPEG-2,4,5-trichloropenylcarbonate, MPEG-p-nitrophenolcarbonate, and various MPEG-succinate derivatives. A number additional polyethylene glycol derivatives and reaction chemistries for attaching polyethylene glycol to proteins are described in WO 98/32466, the entire disclosure of which is incorporated herein by reference. Pegylated protein products produced using the reaction chemistries set out herein are included within the scope of the invention.

**[0127]** The number of polyethylene glycol moieties attached to each protein of the invention (*i.e.*, the degree of substitution) may also vary. For example, the pegylated proteins of the invention may be linked, on average, to 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 17, 20, or more polyethylene glycol molecules. Similarly, the average degree of substitution within ranges such as 1-3, 2-4, 3-5, 4-6, 5-7, 6-8, 7-9, 8-10, 9-11, 10-12, 11-13, 12-14, 13-15, 14-16, 15-17, 16-18, 17-19, or 18-20 polyethylene glycol moieties per protein molecule. Methods for determining the degree of substitution are discussed, for example, in Delgado *et al.*, *Crit. Rev. Thera. Drug Carrier Sys.* 9:249-304 (1992).

**[0128]** The polypeptides of the invention may be in monomers or multimers (*i.e.*, dimers,

trimers, tetramers and higher multimers). Accordingly, the present invention relates to monomers and multimers of the polypeptides of the invention, their preparation, and compositions (preferably, Therapeutics) containing them. In specific embodiments, the polypeptides of the invention are monomers, dimers, trimers or tetramers. In additional embodiments, the multimers of the invention are at least dimers, at least trimers, or at least tetramers.

**[0129]** Multimers encompassed by the invention may be homomers or heteromers. As used herein, the term homomer, refers to a multimer containing only polypeptides corresponding to the amino acid sequence of Table 1 (including fragments, variants, splice variants, and fusion proteins, corresponding to these as described herein). These homomers may contain polypeptides having identical or different amino acid sequences. In a specific embodiment, a homomer of the invention is a multimer containing only polypeptides having an identical amino acid sequence. In another specific embodiment, a homomer of the invention is a multimer containing polypeptides having different amino acid sequences. In specific embodiments, the multimer of the invention is a homodimer (e.g., containing polypeptides having identical or different amino acid sequences) or a homotrimer (e.g., containing polypeptides having identical and/or different amino acid sequences). In additional embodiments, the homomeric multimer of the invention is at least a homodimer, at least a homotrimer, or at least a homotetramer.

**[0130]** As used herein, the term heteromer refers to a multimer containing one or more heterologous polypeptides (i.e., polypeptides of different proteins) in addition to the polypeptides of the invention. In a specific embodiment, the multimer of the invention is a heterodimer, a heterotrimer, or a heterotetramer. In additional embodiments, the heteromeric multimer of the invention is at least a heterodimer, at least a heterotrimer, or at least a heterotetramer.

**[0131]** Multimers of the invention may be the result of hydrophobic, hydrophilic, ionic and/or covalent associations and/or may be indirectly linked, by for example, liposome formation. Thus, in one embodiment, multimers of the invention, such as, for example, homodimers or homotrimers, are formed when polypeptides of the invention contact one another in solution. In another embodiment, heteromultimers of the invention, such as, for example, heterotrimers or heterotetramers, are formed when polypeptides of the invention contact antibodies to the polypeptides of the invention (including antibodies to the heterologous polypeptide sequence in a fusion protein of the invention) in solution. In other

embodiments, multimers of the invention are formed by covalent associations with and/or between the polypeptides of the invention. Such covalent associations may involve one or more amino acid residues contained in the polypeptide sequence (e.g., the polypeptide sequences shown in Table 1). In one instance, the covalent associations are cross-linking between cysteine residues located within the polypeptide sequences which interact in the native (i.e., naturally occurring) polypeptide. In another instance, the covalent associations are the consequence of chemical or recombinant manipulation. Alternatively, such covalent associations may involve one or more amino acid residues contained in the heterologous polypeptide sequence in a fusion protein.

**[0132]** In one example, covalent associations are between the heterologous sequence contained in a fusion protein of the invention (see, e.g., US Patent Number 5,478,925). In a specific example, the covalent associations are between the heterologous sequence contained in a Fc fusion protein of the invention (as described herein). In another specific example, covalent associations of fusion proteins of the invention are between heterologous polypeptide sequence from another protein that is capable of forming covalently associated multimers, such as for example, osteoprotegerin (see, International Publication NO: WO 98/49305, the contents of which is incorporated herein by reference in its entirety). In another embodiment, two or more polypeptides of the invention are joined through peptide linkers. Examples include those peptide linkers described in U.S. Pat. No. 5,073,627 (incorporated herein by reference in its entirety). Proteins comprising multiple polypeptides of the invention separated by peptide linkers may be produced using conventional recombinant DNA technology.

**[0133]** Another method for preparing multimer polypeptides of the invention involves use of polypeptides of the invention fused to a leucine zipper or isoleucine zipper polypeptide sequence. Leucine zipper and isoleucine zipper domains are polypeptides that promote multimerization of the proteins in which they are found. Leucine zippers were originally identified in several DNA-binding proteins (Landschulz et al., Science 240:1759, (1988)), and have since been found in a variety of different proteins. Among the known leucine zippers are naturally occurring peptides and derivatives thereof that dimerize or trimerize. Examples of leucine zipper domains suitable for producing soluble multimeric proteins of the invention are those described in PCT application WO 94/10308, hereby incorporated by reference. Recombinant fusion proteins comprising a polypeptide of the invention fused to a polypeptide sequence that dimerizes or trimerizes in solution are expressed in suitable host cells, and the resulting soluble multimeric fusion protein is recovered from the culture supernatant using

techniques known in the art.

**[0134]** Trimeric polypeptides of the invention may offer the advantage of enhanced biological activity. Preferred leucine zipper moieties and isoleucine moieties are those that preferentially form trimers. One example is a leucine zipper derived from lung surfactant protein D (SPD), as described in Hoppe et al. (FEBS Letters 344:191, (1994)) and in U.S. patent application Ser. No. 08/446,922, hereby incorporated by reference. Other peptides derived from naturally occurring trimeric proteins may be employed in preparing trimeric polypeptides of the invention.

**[0135]** In another example, proteins of the invention are associated by interactions between Flag® polypeptide sequence contained in fusion proteins of the invention containing Flag® polypeptide sequence. In a further embodiment, associations proteins of the invention are associated by interactions between heterologous polypeptide sequence contained in Flag® fusion proteins of the invention and anti-Flag® antibody.

**[0136]** The multimers of the invention may be generated using chemical techniques known in the art. For example, polypeptides desired to be contained in the multimers of the invention may be chemically cross-linked using linker molecules and linker molecule length optimization techniques known in the art (see, e.g., US Patent Number 5,478,925, which is incorporated herein by reference in its entirety). Additionally, multimers of the invention may be generated using techniques known in the art to form one or more inter-molecule cross-links between the cysteine residues located within the sequence of the polypeptides desired to be contained in the multimer (see, e.g., US Patent Number 5,478,925, which is incorporated herein by reference in its entirety). Further, polypeptides of the invention may be routinely modified by the addition of cysteine or biotin to the C-terminus or N-terminus of the polypeptide and techniques known in the art may be applied to generate multimers containing one or more of these modified polypeptides (see, e.g., US Patent Number 5,478,925, which is incorporated herein by reference in its entirety). Additionally, techniques known in the art may be applied to generate liposomes containing the polypeptide components desired to be contained in the multimer of the invention (see, e.g., US Patent Number 5,478,925, which is incorporated herein by reference in its entirety).

**[0137]** Alternatively, multimers of the invention may be generated using genetic engineering techniques known in the art. In one embodiment, polypeptides contained in multimers of the invention are produced recombinantly using fusion protein technology described herein or otherwise known in the art (see, e.g., US Patent Number 5,478,925, which

is incorporated herein by reference in its entirety). In a specific embodiment, polynucleotides coding for a homodimer of the invention are generated by ligating a polynucleotide sequence encoding a polypeptide of the invention to a sequence encoding a linker polypeptide and then further to a synthetic polynucleotide encoding the translated product of the polypeptide in the reverse orientation from the original C-terminus to the N-terminus (lacking the leader sequence) (see, e.g., US Patent Number 5,478,925, which is incorporated herein by reference in its entirety). In another embodiment, recombinant techniques described herein or otherwise known in the art are applied to generate recombinant polypeptides of the invention which contain a transmembrane domain (or hydrophobic or signal peptide) and which can be incorporated by membrane reconstitution techniques into liposomes (see, e.g., US Patent Number 5,478,925, which is incorporated herein by reference in its entirety).

**[0138]     *Polypeptides and Fragments***

**[0139]**     The invention further provides an isolated *S. aureus* polypeptide having an amino acid sequence in Table 1 or SEQ ID NO:1 through 61, or a peptide or polypeptide comprising a portion, fragment, variant or analog of the above polypeptides.

**[0140]     *Variant and Mutant Polypeptides***

**[0141]**     To improve or alter the characteristics of *S. aureus* polypeptides of the present invention, protein engineering may be employed. Recombinant DNA technology known to those skilled in the art can be used to create novel mutant proteins or muteins including single or multiple amino acid substitutions, deletions, additions, or fusion proteins. Such modified polypeptides can show, e.g., increased/decreased activity or increased/decreased stability. In addition, they may be purified in higher yields and show better solubility than the corresponding natural polypeptide, at least under certain purification and storage conditions. Further, the polypeptides of the present invention may be produced as multimers including dimers, trimers and tetramers. Multimerization may be facilitated by linkers or recombinantly though heterologous polypeptides such as Fc regions.

**[0142]     *N-Terminal and C-Terminal Deletion Mutants***

**[0143]**     It is known in the art that one or more amino acids may be deleted from the N-terminus or C-terminus without substantial loss of biological function. For instance, Ron et



al. J. Biol. Chem., 268:2984-2988 (1993), reported modified KGF proteins that had heparin binding activity even if 3, 8, or 27 N-terminal amino acid residues were missing. Accordingly, the present invention provides polypeptides having one or more residues deleted from the amino terminus of the polypeptides shown in Table 1.

**[0144]** Similarly, many examples of biologically functional C-terminal deletion mutants are known. For instance, Interferon gamma shows up to ten times higher activities by deleting 8-10 amino acid residues from the carboxy-terminus of the protein *See, e.g.,* Dobeli, et al. (1988) J. Biotechnology 7:199-216. Accordingly, the present invention provides polypeptides having one or more residues from the carboxy terminus of the polypeptides shown in Table 1. The invention also provides polypeptides having one or more amino acids deleted from both the amino and the carboxyl termini as described below.

**[0145]** The present invention is further directed to polynucleotide encoding portions or fragments of the amino acid sequences described herein as well as to portions or fragments of the isolated amino acid sequences described herein. Fragments include portions of the amino acid sequences of Table 1, at least 7 contiguous amino acid in length, selected from any two integers, one of which representing a N-terminal position. The first codon of the polypeptides of Table 1 is position 1. Every combination of a N-terminal and C-terminal position that a fragment at least 7 contiguous amino acid residues in length could occupy, on any given amino acid sequence of Table 1 is included in the invention. At least means a fragment may be 7 contiguous amino acid residues in length or any integer between 7 and the number of residues in a full length amino acid sequence minus 1. Therefore, included in the invention are contiguous fragments specified by any N-terminal and C-terminal positions of amino acid sequence set forth in Table 1 wherein the contiguous fragment is any integer between 7 and the number of residues in a full length sequence minus 1.

**[0146]** Further, the invention includes polypeptides comprising fragments specified by size, in amino acid residues, rather than by N-terminal and C-terminal positions. The invention includes any fragment size, in contiguous amino acid residues, selected from integers between 7 and the number of residues in a full length sequence minus 1. Preferred sizes of contiguous polypeptide fragments include about 7 amino acid residues, about 10 amino acid residues, about 20 amino acid residues, about 30 amino acid residues, about 40 amino acid residues, about 50 amino acid residues, about 100 amino acid residues, about 200 amino acid residues, about 300 amino acid residues, and about 400 amino acid residues. The preferred sizes are, of course, meant to exemplify, not limit, the present invention as all size fragments representing

any integer between 7 and the number of residues in a full length sequence minus 1 are included in the invention. The present invention also provides for the exclusion of any fragments specified by N-terminal and C-terminal positions or by size in amino acid residues as described above. Any number of fragments specified by N-terminal and C-terminal positions or by size in amino acid residues as described above may be excluded.

**[0147]** The present invention further provides polypeptides having one or more residues deleted from the amino terminus of the amino acid sequence of a polypeptide disclosed herein (e.g., any polypeptide of Table 1). In particular, N-terminal deletions may be described by the general formula m-q, where q is a whole integer representing the total number of amino acid residues in a polypeptide of the invention (e.g., a polypeptide disclosed in Table 1), and m is defined as any integer ranging from 2 to q-6. Polynucleotides encoding these polypeptides are also encompassed by the invention.

**[0148]** The present invention further provides polypeptides having one or more residues from the carboxy-terminus of the amino acid sequence of a polypeptide disclosed herein (e.g., a polypeptide disclosed in Table 1). In particular, C-terminal deletions may be described by the general formula 1-n, where n is any whole integer ranging from 6 to q-1, and where n corresponds to the position of amino acid residue in a polypeptide of the invention. Polynucleotides encoding these polypeptides are also encompassed by the invention.

**[0149]** In addition, any of the above described N- or C-terminal deletions can be combined to produce a N- and C-terminal deleted polypeptide. The invention also provides polypeptides having one or more amino acids deleted from both the amino and the carboxyl termini, which may be described generally as having residues m-n of a polypeptide encoded by a nucleotide sequence (e.g., including, but not limited to the preferred polypeptide disclosed in Table 1), or the cDNA contained in a deposited clone, and/or the complement thereof, where n and m are integers as described above. Polynucleotides encoding these polypeptides are also encompassed by the invention.

**[0150]** The polypeptide fragments of the present invention can be immediately envisaged using the above description and are therefore not individually listed solely for the purpose of not unnecessarily lengthening the specification.

**[0151]** The above fragments need not be active since they would be useful, for example, in immunoassays, in epitope mapping, epitope tagging, to generate antibodies to a particular portion of the polypeptide, as vaccines, and as molecular weight markers.

**[0152]** In addition to N- and C-terminal deletion forms of the protein discussed above, it

also will be recognized by one of ordinary skill in the art that some amino acid sequences of the *S. aureus* polypeptides of the present invention can be varied without significant effect of the structure or function of the protein. If such differences in sequence are contemplated, it should be remembered that there will be critical areas on the protein which determine activity.

[0153] Thus, the invention further includes variations of the *S. aureus* polypeptides which show substantial *S. aureus* polypeptide activity or which include regions of *S. aureus* protein such as the protein portions discussed below. Such mutants include deletions, insertions, inversions, repeats, and substitutions selected according to general rules known in the art so as to have little effect on activity. For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided. There are two main approaches for studying the tolerance of an amino acid sequence to change. See, Bowie, J. U. *et al.* (1990), *Science* 247:1306-1310. The first method relies on the process of evolution, in which mutations are either accepted or rejected by natural selection. The second approach uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene and selections or screens to identify sequences that maintain functionality.

[0154] These studies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The studies indicate which amino acid changes are likely to be permissive at a certain position of the protein. For example, most buried amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Other such phenotypically silent substitutions are described by Bowie *et al.* (*supra*) and the references cited therein. Typically seen as conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu and Ile; interchange of the hydroxyl residues Ser and Thr, exchange of the acidic residues Asp and Glu, substitution between the amide residues Asn and Gln, exchange of the basic residues Lys and Arg and replacements among the aromatic residues Phe, Tyr.

[0155] Thus, the fragment, derivative, analog, or homolog of the polypeptide of Table 1 may be, for example: (i) one in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code; or (ii) one in which one or more of the amino acid residues includes a substituent group; or (iii) one in which the *S. aureus* polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol); or (iv) one in which the additional amino acids are fused to the above form of the polypeptide,

such as an IgG Fc fusion region peptide or leader or secretory sequence or a sequence which is employed for purification of the above form of the polypeptide or a proprotein sequence. Such fragments, derivatives and analogs are deemed to be within the scope of those skilled in the art from the teachings herein.

**[0156]** Thus, the *S. aureus* polypeptides of the present invention may include one or more amino acid substitutions, deletions, or additions, either from natural mutations or human manipulation. As indicated, changes are preferably of a minor nature, such as conservative amino acid substitutions that do not significantly affect the folding or activity of the protein (see Table 3).

**TABLE 3.** Conservative Amino Acid Substitutions.

Aromatic	Phenylalanine Tryptophan Tyrosine
Hydrophobic	Leucine Isoleucine Valine
Polar	Glutamine Asparagine
Basic	Arginine Lysine Histidine
Acidic	Aspartic Acid Glutamic Acid
Small	Alanine Serine Threonine Methionine Glycine

**[0157]** Amino acids in the *S. aureus* proteins of the present invention that are essential for function can be identified by methods known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis. See, e.g., Cunningham et al. (1989) Science 244:1081-1085. The latter procedure introduces single alanine mutations at every residue in the molecule. The resulting mutant molecules are then tested for biological activity using assays appropriate for measuring the function of the particular protein.

**[0158]** Of special interest are substitutions of charged amino acids with other charged or neutral amino acids which may produce proteins with highly desirable improved characteristics, such as less aggregation. Aggregation may not only reduce activity but also be

problematic when preparing pharmaceutical formulations, because aggregates can be immunogenic. See, e.g., Pinckard et al., (1967) Clin. Exp. Immunol. 2:331-340; Robbins, et al., (1987) Diabetes 36:838-845; Cleland, et al., (1993) Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377.

**[0159]** The polypeptides of the present invention are preferably provided in an isolated form, and may partially or substantially purified. A recombinantly produced version of the *S. aureus* polypeptide can be substantially purified by the one-step method described by Smith et al. (1988) Gene 67:31-40. Polypeptides of the invention also can be purified from natural or recombinant sources using antibodies directed against the polypeptides of the invention in methods which are well known in the art of protein purification. The purity of the polypeptide of the present invention may also specified in percent purity as relative to heterologous containing polypeptides. Preferred purities include at least 25%, 50%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.5%, 99.75%, and 100% pure, as relative to heretologous containing polypeptides.

**[0160]** The invention provides for isolated *S. aureus* polypeptides comprising an the amino acid sequence of a full-length *S. aureus* polypeptide having the complete amino acid sequence shown in Table 1 and the amino acid sequence of a full-length *S. aureus* polypeptide having the complete amino acid sequence shown in Table 1 excepting the N-terminal codon (e.g. including, but not limited to, methionine, leucine, and/or valine) The polypeptides of the present invention also include polypeptides having an amino acid sequence at least 80% identical, more preferably at least 90% identical, and still more preferably 95%, 96%, 97%, 98% or 99% identical to a member of the group consisting of (a) a polypeptide encoded by any of the polynucleotide sequences shown in Table 1, (b) any of the polypeptide sequences shown in Table 1 and (c) the complement of a polynucleotide sequence encoding the polypeptide of (a) or (b) above. Further polypeptides of the present invention include polypeptides which have at least 90% similarity, more preferably at least 95% similarity, and still more preferably at least 96%, 97%, 98% or 99% similarity to those described above.

**[0161]** A further embodiment of the invention relates to a polypeptide which comprises the amino acid sequence of a *S. aureus* polypeptide having an amino acid sequence which contains at least one conservative amino acid substitution, but not more than 50 conservative amino acid substitutions, not more than 40 conservative amino acid substitutions, not more than 30 conservative amino acid substitutions, and not more than 20 conservative amino acid substitutions. Also provided are polypeptides which comprise the amino acid sequence of a *S.*

*aureus* polypeptide, having at least one, but not more than 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 conservative amino acid substitutions.

**[0162]** By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence except that the subject polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query amino acid sequence, up to 5% (5 of 100) of the amino acid residues in the subject sequence may be inserted, deleted, (indels) or substituted with another amino acid. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

**[0163]** As a practical matter, whether any particular polypeptide is at least 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequences shown in Table 1, or a fragment thereof can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al., (1990) Comp. App. Biosci. 6:237-245. In a sequence alignment the query and subject sequences are both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter.

**[0164]** If the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, the results, in percent identity, must be manually corrected. This is because the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding

subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the purposes of manually adjusting the percent identity score. That is, only query amino acid residues outside the farthest N- and C-terminal residues of the subject sequence.

[0165] For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not match/align with the first 10 residues at the N-terminus. The 10 unpaired residues represent 10% of the sequence (number of residues at the N- and C- termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence. This time the deletions are internal so there are no residues at the N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query sequence are manually corrected. No other manual corrections are to be made for the purposes of the present invention.

[0166] The above polypeptide sequences are included irrespective of whether they have their normal biological activity. This is because even where a particular polypeptide molecule does not have biological activity, one of skill in the art would still know how to use the polypeptide, for instance, as a vaccine or to generate antibodies. Other uses of the polypeptides of the present invention that do not have *S. aureus* activity include, *inter alia*, as epitope tags, in epitope mapping, and as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods known to those of skill in the art.

[0167] As described below, the polypeptides of the present invention can also be used to raise polyclonal and monoclonal antibodies, which are useful in assays for detecting *S. aureus* protein expression or as agonists and antagonists capable of enhancing or inhibiting *S. aureus*

protein function. Further, such polypeptides can be used in the yeast two-hybrid system to "capture" *S. aureus* protein binding proteins which are also candidate agonists and antagonists according to the present invention. *See, e.g.,* Fields et al. (1989) Nature 340:245-246.

**[0168]     *Epitopes and Antibodies***

**[0169]**     The present invention encompasses polypeptides comprising, or alternatively consisting of, an epitope of the polypeptide having an amino acid sequence in Table 1, or encoded by a polynucleotide that hybridizes to the complement of a nucleotide sequence shown in Table 1 under stringent hybridization conditions or alternatively, lower stringency hybridization conditions as defined supra. The present invention further encompasses polynucleotide sequences encoding an epitope of a polypeptide sequence of the invention (such as, for example, a nucleotide sequence disclosed in Table 1), polynucleotide sequences of the complementary strand of a polynucleotide sequence encoding an epitope of the invention, and polynucleotide sequences which hybridize to the complementary strand under stringent hybridization conditions or lower stringency hybridization conditions defined supra.

**[0170]**     The term "epitopes," as used herein, refers to portions of a polypeptide having antigenic or immunogenic activity in an animal, preferably a mammal, and most preferably in a human. In a preferred embodiment, the present invention encompasses a polypeptide comprising an epitope, as well as the polynucleotide encoding this polypeptide. An "immunogenic epitope," as used herein, is defined as a portion of a protein that elicits an antibody response in an animal, as determined by any method known in the art, for example, by the methods for generating antibodies described infra. (See, for example, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998- 4002 (1983)). The term "antigenic epitope," as used herein, is defined as a portion of a protein to which an antibody can immunospecifically bind its antigen as determined by any method well known in the art, for example, by the immunoassays described herein. Immunospecific binding excludes non-specific binding but does not necessarily exclude cross-reactivity with other antigens. Antigenic epitopes need not necessarily be immunogenic.

**[0171]**     Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985), further described in U.S. Patent No. 4,631,211).

**[0172]**     In the present invention, antigenic epitopes preferably contain a sequence of at least 4, at least 5, at least 6, at least 7, more preferably at least 8, at least 9, at least 10, at least 11, at



least 12, at least 13, at least 14, at least 15, at least 20, at least 25, at least 30, at least 40, at least 50, and, most preferably, between about 15 to about 30 amino acids. Preferred polypeptides comprising immunogenic or antigenic epitopes are at least 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 amino acid residues in length. Additional non-exclusive preferred antigenic epitopes include the antigenic epitopes disclosed herein, as well as portions thereof. Antigenic epitopes are useful, for example, to raise antibodies, including monoclonal antibodies, that specifically bind the epitope. Preferred antigenic epitopes include the antigenic epitopes disclosed herein, as well as any combination of two, three, four, five or more of these antigenic epitopes. Antigenic epitopes can be used as the target molecules in immunoassays. (See, for instance, Wilson et al., *Cell* 37:767-778 (1984); Sutcliffe et al., *Science* 219:660-666 (1983)).

**[0173]** Similarly, immunogenic epitopes can be used, for example, to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., *supra*; Wilson et al., *supra*; Chow et al., *Proc. Natl. Acad. Sci. USA* 82:910-914; and Bittle et al., *J. Gen. Virol.* 66:2347-2354 (1985). Preferred immunogenic epitopes include the immunogenic epitopes disclosed herein, as well as any combination of two, three, four, five or more of these immunogenic epitopes. The polypeptides comprising one or more immunogenic epitopes may be presented for eliciting an antibody response together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse), or, if the polypeptide is of sufficient length (at least about 25 amino acids), the polypeptide may be presented without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a denatured polypeptide (e.g., in Western blotting).

**[0174]** Epitope-bearing polypeptides of the present invention may be used to induce antibodies according to methods well known in the art including, but not limited to, in vivo immunization, in vitro immunization, and phage display methods. See, e.g., Sutcliffe et al., *supra*; Wilson et al., *supra*, and Bittle et al., *J. Gen. Virol.*, 66:2347-2354 (1985). If in vivo immunization is used, animals may be immunized with free peptide; however, anti-peptide antibody titer may be boosted by coupling the peptide to a macromolecular carrier, such as keyhole limpet hemacyanin (KLH) or tetanus toxoid. For instance, peptides containing cysteine residues may be coupled to a carrier using a linker such as maleimidobenzoyl-N-hydroxysuccinimide ester (MBS), while other peptides may be coupled to carriers using a more general linking agent such as glutaraldehyde. Animals such as rabbits, rats and mice are

immunized with either free or carrier- coupled peptides, for instance, by intraperitoneal and/or intradermal injection of emulsions containing about 100 µg of peptide or carrier protein and Freund's adjuvant or any other adjuvant known for stimulating an immune response. Several booster injections may be needed, for instance, at intervals of about two weeks, to provide a useful titer of anti-peptide antibody which can be detected, for example, by ELISA assay using free peptide adsorbed to a solid surface. The titer of anti-peptide antibodies in serum from an immunized animal may be increased by selection of anti-peptide antibodies, for instance, by adsorption to the peptide on a solid support and elution of the selected antibodies according to methods well known in the art.

**[0175]** As one of skill in the art will appreciate, and as discussed above, the polypeptides of the present invention comprising an immunogenic or antigenic epitope can be fused to other polypeptide sequences. For example, the polypeptides of the present invention may be fused with the constant domain of immunoglobulins (IgA, IgE, IgG, IgM), or portions thereof (CH1, CH2, CH3, or any combination thereof and portions thereof) resulting in chimeric polypeptides. Such fusion proteins may facilitate purification and may increase half-life in vivo. This has been shown for chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. See, e.g., EP 394,827; Traunecker et al., *Nature*, 331:84-86 (1988). Enhanced delivery of an antigen across the epithelial barrier to the immune system has been demonstrated for antigens (e.g., insulin) conjugated to an FcRn binding partner such as IgG or Fc fragments (see, e.g., PCT Publications WO 96/22024 and WO 99/04813). IgG Fusion proteins that have a disulfide-linked dimeric structure due to the IgG portion disulfide bonds have also been found to be more efficient in binding and neutralizing other molecules than monomeric polypeptides or fragments thereof alone. See, e.g., Fountoulakis et al., *J. Biochem.*, 270:3958-3964 (1995). Nucleic acids encoding the above epitopes can also be recombined with a gene of interest as an epitope tag (e.g., the hemagglutinin ("HA") tag or flag tag) to aid in detection and purification of the expressed polypeptide. For example, a system described by Janknecht et al. allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (Janknecht et al., 1991, *Proc. Natl. Acad. Sci. USA* 88:8972- 897). In this system, the gene of interest is subcloned into a vaccinia recombination plasmid such that the open reading frame of the gene is translationally fused to an amino-terminal tag consisting of six histidine residues. The tag serves as a matrix binding domain for the fusion protein. Extracts from cells infected with the

recombinant vaccinia virus are loaded onto Ni<sup>2+</sup> nitriloacetic acid-agarose column and histidine-tagged proteins can be selectively eluted with imidazole-containing buffers.

**[0176]** Additional fusion proteins of the invention may be generated through the techniques of gene-shuffling, motif-shuffling, exon-shuffling, and/or codon-shuffling (collectively referred to as "DNA shuffling"). DNA shuffling may be employed to modulate the activities of polypeptides of the invention, such methods can be used to generate polypeptides with altered activity, as well as agonists and antagonists of the polypeptides. See, generally, U.S. Patent Nos. 5,605,793; 5,811,238; 5,830,721; 5,834,252; and 5,837,458, and Patten et al., Curr. Opin. Biotechnol. 8:724-33 (1997); Harayama, Trends Biotechnol. 16(2):76-82 (1998); Hansson, et al., J. Mol. Biol. 287:265-76 (1999); and Lorenzo and Blasco, Biotechniques 24(2):308- 13 (1998) (each of these patents and publications are incorporated herein by reference in its entirety). In one embodiment, alteration of polynucleotides corresponding to those shown in Table 1 and the polypeptides encoded by these polynucleotides may be achieved by DNA shuffling. DNA shuffling involves the assembly of two or more DNA segments by homologous or site-specific recombination to generate variation in the polynucleotide sequence. In another embodiment, polynucleotides of the invention, or the encoded polypeptides, may be altered by being subjected to random mutagenesis by error-prone PCR, random nucleotide insertion or other methods prior to recombination. In another embodiment, one or more components, motifs, sections, parts, domains, fragments, etc., of a polynucleotide encoding a polypeptide of the invention may be recombined with one or more components, motifs, sections, parts, domains, fragments, etc. of one or more heterologous molecules.

**[0177]** Predicted antigenic epitopes are shown in Table 4, below. It is pointed out that Table 4 only lists amino acid residues comprising epitopes predicted to have the highest degree of antigenicity by a particular algorithm. The polypeptides not listed in Table 4 and portions of polypeptides not listed in Table 4 are not considered non-antigenic. This is because they may still be antigenic *in vivo* but merely not recognized as such by the particular algorithm used. Thus, Table 4 lists the amino acids residues comprising only preferred antigenic epitopes, not a complete list. In fact, all fragments of the polypeptide sequence of Table 1, at least 7 amino acid residues in length, are included in the present invention as being useful in epitope mapping and in making antibodies to particular portions of the polypeptides. Moreover, Table 4 lists only the critical residues of the epitopes determined by the Jameson-Wolf analysis. Thus, additional flanking residues on either the N-terminal, C-terminal, or both N- and C-

terminal ends may be added to the sequences of Table 4 to generate an epitope-bearing protein a least 7 residues in length. Amino acid residues comprising other antigenic epitopes may be determined by algorithms similar to the Jameson-Wolf analysis or by *in vivo* testing for an antigenic response using the methods described herein or those known in the art.

**TABLE 4. Residues Comprising Antigenic Epitopes**

HGS001	from about Asp-47 to about Asp-50, from about Ser-128 to about Asp-130, from about Lys-265 to about Gly-267.
HGS005	from about Arg-104 to about Asp-106, from about Lys-116 to about Lys-120.
HGS007m	from about Glu-155 to about Gly-158, from about Gln-178 to about Gly-181, from about Ser-304 to about Cys-306, from about Asp-401 to about Tyr-403, from about Asn-405 to about Gly-408, from about Asp-411 to about Gly-416.
HGS009	from about Pro-257 to about Lys-259.
HGS014	from about Arg-186 to about Asp-188.
HGS019	from about Lys-98 to about Gly-100, from about Pro-187 to about Asp-189.
HGS023	from about Ser-251 to about Gly-253, from about Lys-437 to about Lys-440.
HGS025	from about Met-51 to about Gly-53.
HGS026	from about Asn-105 to about Lys-108, from about Glu-190 to about Gly-193, from about Arg-226 to about Ala-230.
HGS028	from about Ile-10 to about Tyr-13.
HGS030	from about Glu-11 to about Gly-14, from about Arg-147 to about Gln-149.
HGS033	from about Lys-143 to about Ser-145.
HGS034	from about Pro-33 to about Gln-35.
HGS036	from about Asp-64 to about Tyr-66, from about Asp-255 to about Tyr-257.
HGS040	from about Pro-30 to about Lys-32, from about Asp-76 to about Asp-78.
168153_3	from about Asn-35 to about Arg-37, from about Pro-135 to about Asp-138, from about Pro-185 to about Gln-188.
168153_2	from about Asp-54 to about Arg-56.
168153_1	from about Lys-64 to about Asp-67, from about Gln-319 to about Lys-322, from about Asn-342 to about Lys-344.
168339_2	from about Asn-82 to about Arg-85.

[0178] Non-limiting examples of antigenic polypeptides or peptides that can be used to generate an Staphylococcal-specific immune response or antibodies include fragments of the amino acid sequences of Table 1 as discussed above. Table 4 discloses a list of non-limiting residues that are involved in the antigenicity of the epitope-bearing fragments of the present invention. Therefore, also included in the present inventions are isolated and purified antigenic epitope-bearing fragments of the polypeptides of the present invention comprising a peptide sequences of Table 4. The antigenic epitope-bearing fragments comprising a peptide sequence of Table 4 preferably contain between 7 to 50 amino acids (i.e. any integer between 7 and 50) of a polypeptide of the present invention. Also, included in the present invention are antigenic polypeptides between the integers of 7 and the full length sequence of a polypeptide of Table 1

comprising 1 or more amino acid sequences of Table 4. Therefore, in most cases, the polypeptides of Table 4 make up only a portion of the antigenic polypeptide. All combinations of sequences between the integers of 7 and the full sequence of a polypeptide sequence of Table 1 are included. The antigenic epitope-bearing fragments may be specified by either the number of contiguous amino acid residues or by specific N-terminal and C-terminal positions as described above for the polypeptide fragments of the present invention, wherein the first codon of each polypeptide sequence of Table 1 is position 1. Any number of the described antigenic epitope-bearing fragments of the present invention may also be excluded from the present invention in the same manner.

**[0179]     Antibodies**

**[0180]**     Further polypeptides of the invention relate to antibodies and T-cell antigen receptors (TCR) which immunospecifically bind a polypeptide, polypeptide fragment, or variant of a polypeptide sequence shown in Table 1, and/or an epitope, of the present invention (as determined by immunoassays well known in the art for assaying specific antibody-antigen binding). Antibodies of the invention include, but are not limited to, polyclonal, monoclonal, multispecific, human, humanized or chimeric antibodies, single chain antibodies, Fab fragments, F(ab') fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies (including, e.g., anti-Id antibodies to antibodies of the invention), and epitope-binding fragments of any of the above. The term "antibody," as used herein, refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, i.e., molecules that contain an antigen binding site that immunospecifically binds an antigen. The immunoglobulin molecules of the invention can be of any type (e.g., IgG, IgE, IgM, IgD, IgA and IgY), class (e.g., IgG1, IgG2, IgG3, IgG4, IgA1 and IgA2) or subclass of immunoglobulin molecule.

**[0181]**     Most preferably the antibodies are human antigen-binding antibody fragments of the present invention and include, but are not limited to, Fab, Fab' and F(ab')<sub>2</sub>, Fd, single-chain Fvs (scFv), single-chain antibodies, disulfide-linked Fvs (sdFv) and fragments comprising either a VL or VH domain. Antigen-binding antibody fragments, including single-chain antibodies, may comprise the variable region(s) alone or in combination with the entirety or a portion of the following: hinge region, CH1, CH2, and CH3 domains. Also included in the invention are antigen-binding fragments also comprising any combination of variable region(s) with a hinge region, CH1, CH2, and CH3 domains. The antibodies of the invention

may be from any animal origin including birds and mammals. Preferably, the antibodies are human, murine (e.g., mouse and rat), donkey, ship rabbit, goat, guinea pig, camel, horse, or chicken. As used herein, "human" antibodies include antibodies having the amino acid sequence of a human immunoglobulin and include antibodies isolated from human immunoglobulin libraries or from animals transgenic for one or more human immunoglobulin and that do not express endogenous immunoglobulins, as described infra and, for example in, U.S. Patent No. 5,939,598 by Kucherlapati et al.

**[0182]** The antibodies of the present invention may be monospecific, bispecific, trispecific or of greater multispecificity. Multispecific antibodies may be specific for different epitopes of a polypeptide of the present invention or may be specific for both a polypeptide of the present invention as well as for a heterologous epitope, such as a heterologous polypeptide or solid support material. See, e.g., PCT publications WO 93/17715; WO 92/08802; WO 91/00360; WO 92/05793; Tutt, et al., J. Immunol. 147:60-69 (1991); U.S. Patent Nos. 4,474,893; 4,714,681; 4,925,648; 5,573,920; 5,601,819; Kostelny et al., J. Immunol. 148:1547-1553 (1992).

**[0183]** Antibodies of the present invention may be described or specified in terms of the epitope(s) or portion(s) of a polypeptide of the present invention which they recognize or specifically bind. The epitope(s) or polypeptide portion(s) may be specified as described herein, e.g., by N-terminal and C-terminal positions, by size in contiguous amino acid residues, or listed in the Tables and sequence listing. Antibodies which specifically bind any epitope or polypeptide of the present invention may also be excluded. Therefore, the present invention includes antibodies that specifically bind polypeptides of the present invention, and allows for the exclusion of the same.

**[0184]** Antibodies of the present invention may also be described or specified in terms of their cross-reactivity. Antibodies that do not bind any other analog, ortholog, or homolog of a polypeptide of the present invention are included. Antibodies that bind polypeptides with at least 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 70%, at least 65%, at least 60%, at least 55%, and at least 50% identity (as calculated using methods known in the art and described herein) to a polypeptide of the present invention are also included in the present invention. In specific embodiments, antibodies of the present invention cross-react with murine, rat and/or rabbit homologs of human proteins and the corresponding epitopes thereof. Antibodies that do not bind polypeptides with less than 95%, less than 90%, less than 85%, less than 80%, less than 75%, less than 70%, less than 65%, less than 60%, less than

55%, and less than 50% identity (as calculated using methods known in the art and described herein) to a polypeptide of the present invention are also included in the present invention. In a specific embodiment, the above-described cross-reactivity is with respect to any single specific antigenic or immunogenic polypeptide, or combination(s) of 2, 3, 4, 5, or more of the specific antigenic and/or immunogenic polypeptides disclosed herein. Further included in the present invention are antibodies which bind polypeptides encoded by polynucleotides which hybridize to a polynucleotide of the present invention under stringent hybridization conditions (as described herein). Antibodies of the present invention may also be described or specified in terms of their binding affinity to a polypeptide of the invention. Preferred binding affinities include those with a dissociation constant or  $K_d$  less than  $5 \times 10^{-2}$  M,  $10^{-2}$  M,  $5 \times 10^{-3}$  M,  $10^{-3}$  M,  $5 \times 10^{-4}$  M,  $10^{-4}$  M,  $5 \times 10^{-5}$  M,  $10^{-5}$  M,  $5 \times 10^{-6}$  M,  $10^{-6}$  M,  $5 \times 10^{-7}$  M,  $10^{-7}$  M,  $5 \times 10^{-8}$  M,  $10^{-8}$  M,  $5 \times 10^{-9}$  M,  $10^{-9}$  M,  $5 \times 10^{-10}$  M,  $10^{-10}$  M,  $5 \times 10^{-11}$  M,  $10^{-11}$  M,  $5 \times 10^{-12}$  M,  $10^{-12}$  M,  $5 \times 10^{-13}$  M,  $10^{-13}$  M,  $5 \times 10^{-14}$  M,  $10^{-14}$  M,  $5 \times 10^{-15}$  M, or  $10^{-15}$  M.

**[0185]** The invention also provides antibodies that competitively inhibit binding of an antibody to an epitope of the invention as determined by any method known in the art for determining competitive binding, for example, the immunoassays described herein. In preferred embodiments, the antibody competitively inhibits binding to the epitope by at least 95%, at least 90%, at least 85 %, at least 80%, at least 75%, at least 70%, at least 60%, or at least 50%.

**[0186]** Antibodies of the present invention may act as agonists or antagonists of the polypeptides of the present invention. For example, the present invention includes antibodies which disrupt the receptor/ligand interactions with the polypeptides of the invention either partially or fully. Preferably, antibodies of the present invention bind an antigenic epitope disclosed herein, or a portion thereof. The invention features both receptor-specific antibodies and ligand-specific antibodies. The invention also features receptor-specific antibodies which do not prevent ligand binding but prevent receptor activation. Receptor activation (i.e., signaling) may be determined by techniques described herein or otherwise known in the art. For example, receptor activation can be determined by detecting the phosphorylation (e.g., tyrosine or serine/threonine) of the receptor or its substrate by immunoprecipitation followed by western blot analysis (for example, as described supra). In specific embodiments, antibodies are provided that inhibit ligand activity or receptor activity by at least 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 70%, at least 60%, or at least 50% of the activity in absence of the antibody.

**[0187]** The invention also features receptor-specific antibodies which both prevent ligand binding and receptor activation as well as antibodies that recognize the receptor-ligand complex, and, preferably, do not specifically recognize the unbound receptor or the unbound ligand. Likewise, included in the invention are neutralizing antibodies which bind the ligand and prevent binding of the ligand to the receptor, as well as antibodies which bind the ligand, thereby preventing receptor activation, but do not prevent the ligand from binding the receptor. Further included in the invention are antibodies which activate the receptor. These antibodies may act as receptor agonists, i.e., potentiate or activate either all or a subset of the biological activities of the ligand-mediated receptor activation, for example, by inducing dimerization of the receptor. The antibodies may be specified as agonists, antagonists or inverse agonists for biological activities comprising the specific biological activities of the peptides of the invention disclosed herein. The above antibody agonists can be made using methods known in the art. See, e.g., PCT publication WO 96/40281; U.S. Patent No. 5,811,097; Deng et al., *Blood* 92(6):1981-1988 (1998); Chen et al., *Cancer Res.* 58(16):3668-3678 (1998); Harrop et al., *J. Immunol.* 161(4):1786-1794 (1998); Zhu et al., *Cancer Res.* 58(15):3209-3214 (1998); Yoon et al., *J. Immunol.* 160(7):3170-3179 (1998); Prat et al., *J. Cell. Sci.* 111(Pt2):237-247 (1998); Pitard et al., *J. Immunol. Methods* 205(2):177-190 (1997); Liautard et al., *Cytokine* 9(4):233-241 (1997); Carlson et al., *J. Biol. Chem.* 272(17):11295-11301 (1997); Taryman et al., *Neuron* 14(4):755-762 (1995); Muller et al., *Structure* 6(9):1153-1167 (1998); Bartunek et al., *Cytokine* 8(1):14-20 (1996) (which are all incorporated by reference herein in their entireties).

**[0188]** Antibodies of the present invention may be used, for example, but not limited to, to purify, detect, and target the polypeptides of the present invention, including both in vitro and in vivo diagnostic and therapeutic methods. For example, the antibodies have use in immunoassays for qualitatively and quantitatively measuring levels of the polypeptides of the present invention in biological samples. See, e.g., Harlow et al., *Antibodies: A Laboratory Manual*, (Cold Spring Harbor Laboratory Press, 2nd ed. 1988) (incorporated by reference herein in its entirety).

**[0189]** As discussed in more detail below, the antibodies of the present invention may be used either alone or in combination with other compositions. The antibodies may further be recombinantly fused to a heterologous polypeptide at the N- or C-terminus or chemically conjugated (including covalently and non-covalently conjugations) to polypeptides or other compositions. For example, antibodies of the present invention may be recombinantly fused



or conjugated to molecules useful as labels in detection assays and effector molecules such as heterologous polypeptides, drugs, radionuclides, or toxins. See, e.g., PCT publications WO 92/08495; WO 91/14438; WO 89/12624; U.S. Patent No. 5,314,995; and EP 396,387.

**[0190]** The antibodies of the invention include derivatives that are modified, i.e., by the covalent attachment of any type of molecule to the antibody such that covalent attachment does not prevent the antibody from generating an anti-idiotypic response. For example, but not by way of limitation, the antibody derivatives include antibodies that have been modified, e.g., by glycosylation, acetylation, pegylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to a cellular ligand or other protein, etc. Any of numerous chemical modifications may be carried out by known techniques, including, but not limited to specific chemical cleavage, acetylation, formylation, metabolic synthesis of tunicamycin, etc. Additionally, the derivative may contain one or more non-classical amino acids.

**[0191]** The antibodies of the present invention may be generated by any suitable method known in the art. Polyclonal antibodies to an antigen-of-interest can be produced by various procedures well known in the art. For example, a polypeptide of the invention can be administered to various host animals including, but not limited to, rabbits, mice, rats, etc. to induce the production of sera containing polyclonal antibodies specific for the antigen. Various adjuvants may be used to increase the immunological response, depending on the host species, and include but are not limited to, Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanins, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and corynebacterium parvum. Such adjuvants are also well known in the art.

**[0192]** Monoclonal antibodies can be prepared using a wide variety of techniques known in the art including the use of hybridoma, recombinant, and phage display technologies, or a combination thereof. For example, monoclonal antibodies can be produced using hybridoma techniques including those known in the art and taught, for example, in Harlow et al., *Antibodies: A Laboratory Manual*, (Cold Spring Harbor Laboratory Press, 2nd ed. 1988); Hammerling, et al., in: *Monoclonal Antibodies and T-Cell Hybridomas* 563-681 (Elsevier, N.Y., 1981) (said references incorporated by reference in their entireties). The term "monoclonal antibody" as used herein is not limited to antibodies produced through hybridoma technology. The term "monoclonal antibody" refers to an antibody that is derived from a

single clone, including any eukaryotic, prokaryotic, or phage clone, and not the method by which it is produced.

**[0193]** Methods for producing and screening for specific antibodies using hybridoma technology are routine and well known in the art and are discussed in detail in the Examples (e.g., Example 16). In a non-limiting example, mice can be immunized with a polypeptide of the invention or a cell expressing such peptide. Once an immune response is detected, e.g., antibodies specific for the antigen are detected in the mouse serum, the mouse spleen is harvested and splenocytes isolated. The splenocytes are then fused by well known techniques to any suitable myeloma cells, for example cells from cell line SP20 available from the ATCC. Hybridomas are selected and cloned by limited dilution. The hybridoma clones are then assayed by methods known in the art for cells that secrete antibodies capable of binding a polypeptide of the invention. Ascites fluid, which generally contains high levels of antibodies, can be generated by immunizing mice with positive hybridoma clones.

**[0194]** Accordingly, the present invention provides methods of generating monoclonal antibodies as well as antibodies produced by the method comprising culturing a hybridoma cell secreting an antibody of the invention wherein, preferably, the hybridoma is generated by fusing splenocytes isolated from a mouse immunized with an antigen of the invention with myeloma cells and then screening the hybridomas resulting from the fusion for hybridoma clones that secrete an antibody able to bind a polypeptide of the invention.

**[0195]** Antibody fragments which recognize specific epitopes may be generated by known techniques. For example, Fab and F(ab')<sub>2</sub> fragments of the invention may be produced by proteolytic cleavage of immunoglobulin molecules, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')<sub>2</sub> fragments). F(ab')<sub>2</sub> fragments contain the variable region, the light chain constant region and the CH1 domain of the heavy chain.

**[0196]** For example, the antibodies of the present invention can also be generated using various phage display methods known in the art. In phage display methods, functional antibody domains are displayed on the surface of phage particles which carry the polynucleotide sequences encoding them. In a particular embodiment, such phage can be utilized to display antigen binding domains expressed from a repertoire or combinatorial antibody library (e.g., human or murine). Phage expressing an antigen binding domain that binds the antigen of interest can be selected or identified with antigen, e.g., using labeled antigen or antigen bound or captured to a solid surface or bead. Phage used in these methods are typically filamentous phage including fd and M13 binding domains expressed from phage

with Fab, Fv or disulfide stabilized Fv antibody domains recombinantly fused to either the phage gene III or gene VIII protein. Examples of phage display methods that can be used to make the antibodies of the present invention include those disclosed in Brinkman et al., J. Immunol. Methods 182:41-50 (1995); Ames et al., J. Immunol. Methods 184:177-186 (1995); Kettleborough et al., Eur. J. Immunol. 24:952-958 (1994); Persic et al., Gene 187 9-18 (1997); Burton et al., Advances in Immunology 57:191-280 (1994); PCT application No. PCT/GB91/01134; PCT publications WO 90/02809; WO 91/10737; WO 92/01047; WO 92/18619; WO 93/11236; WO 95/15982; WO 95/20401; and U.S. Patent Nos. 5,698,426; 5,223,409; 5,403,484; 5,580,717; 5,427,908; 5,750,753; 5,821,047; 5,571,698; 5,427,908; 5,516,637; 5,780,225; 5,658,727; 5,733,743 and 5,969,108; each of which is incorporated herein by reference in its entirety.

**[0197]** As described in the above references, after phage selection, the antibody coding regions from the phage can be isolated and used to generate whole antibodies, including human antibodies, or any other desired antigen binding fragment, and expressed in any desired host, including mammalian cells, insect cells, plant cells, yeast, and bacteria, e.g., as described in detail below. For example, techniques to recombinantly produce Fab, Fab' and F(ab')<sub>2</sub> fragments can also be employed using methods known in the art such as those disclosed in PCT publication WO 92/22324; Mullinax et al., BioTechniques 12(6):864-869 (1992); and Sawai et al., AJRI 34:26-34 (1995); and Better et al., Science 240:1041-1043 (1988) (said references incorporated by reference in their entireties).

**[0198]** Examples of techniques which can be used to produce single-chain Fvs and antibodies include those described in U.S. Patents 4,946,778 and 5,258,498; Huston et al., Methods in Enzymology 203:46-88 (1991); Shu et al., PNAS 90:7995-7999 (1993); and Skerra et al., Science 240:1038-1040 (1988). For some uses, including in vivo use of antibodies in humans and in vitro detection assays, it may be preferable to use chimeric, humanized, or human antibodies. A chimeric antibody is a molecule in which different portions of the antibody are derived from different animal species, such as antibodies having a variable region derived from a murine monoclonal antibody and a human immunoglobulin constant region. Methods for producing chimeric antibodies are known in the art. See e.g., Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Gillies et al., (1989) J. Immunol. Methods 125:191-202; U.S. Patent Nos. 5,807,715; 4,816,567; and 4,816,397, which are incorporated herein by reference in their entirety. Humanized antibodies are antibody molecules from non-human species antibody that binds the desired antigen having

one or more complementarity determining regions (CDRs) from the non-human species and a framework regions from a human immunoglobulin molecule. Often, framework residues in the human framework regions will be substituted with the corresponding residue from the CDR donor antibody to alter, preferably improve, antigen binding. These framework substitutions are identified by methods well known in the art, e.g., by modeling of the interactions of the CDR and framework residues to identify framework residues important for antigen binding and sequence comparison to identify unusual framework residues at particular positions. (See, e.g., Queen et al., U.S. Patent No. 5,585,089; Riechmann et al., *Nature* 332:323 (1988), which are incorporated herein by reference in their entireties.) Antibodies can be humanized using a variety of techniques known in the art including, for example, CDR-grafting (EP 239,400; PCT publication WO 91/09967; U.S. Patent Nos. 5,225,539; 5,530,101; and 5,585,089), veneering or resurfacing (EP 592,106; EP 519,596; Padlan, *Molecular Immunology* 28(4/5):489-498 (1991); Studnicka et al., *Protein Engineering* 7(6):805-814 (1994); Roguska. et al., *PNAS* 91:969-973 (1994)), and chain shuffling (U.S. Patent No. 5,565,332).

**[0199]** Completely human antibodies are particularly desirable for therapeutic treatment of human patients. Human antibodies can be made by a variety of methods known in the art including phage display methods described above using antibody libraries derived from human immunoglobulin sequences. See also, U.S. Patent Nos. 4,444,887 and 4,716,111; and PCT publications WO 98/46645, WO 98/50433, WO 98/24893, WO 98/16654, WO 96/34096, WO 96/33735, and WO 91/10741; each of which is incorporated herein by reference in its entirety.

**[0200]** Human antibodies can also be produced using transgenic mice which are incapable of expressing functional endogenous immunoglobulins, but which can express human immunoglobulin genes. For example, the human heavy and light chain immunoglobulin gene complexes may be introduced randomly or by homologous recombination into mouse embryonic stem cells. Alternatively, the human variable region, constant region, and diversity region may be introduced into mouse embryonic stem cells in addition to the human heavy and light chain genes. The mouse heavy and light chain immunoglobulin genes may be rendered non-functional separately or simultaneously with the introduction of human immunoglobulin loci by homologous recombination. In particular, homozygous deletion of the JH region prevents endogenous antibody production. The modified embryonic stem cells are expanded and microinjected into blastocysts to produce chimeric mice. The chimeric mice are then bred to produce homozygous offspring which express human antibodies. The transgenic mice are immunized in the normal fashion with a selected antigen, e.g., all or a portion of a polypeptide

of the invention. Monoclonal antibodies directed against the antigen can be obtained from the immunized, transgenic mice using conventional hybridoma technology. The human immunoglobulin transgenes harbored by the transgenic mice rearrange during B cell differentiation, and subsequently undergo class switching and somatic mutation. Thus, using such a technique, it is possible to produce therapeutically useful IgG, IgA, IgM and IgE antibodies. For an overview of this technology for producing human antibodies, see Lonberg and Huszar, *Int. Rev. Immunol.* 13:65-93 (1995). For a detailed discussion of this technology for producing human antibodies and human monoclonal antibodies and protocols for producing such antibodies, see, e.g., PCT publications WO 98/24893; WO 92/01047; WO 96/34096; WO 96/33735; European Patent No. 0 598 877; U.S. Patent Nos. 5,413,923; 5,625,126; 5,633,425; 5,569,825; 5,661,016; 5,545,806; 5,814,318; 5,885,793; 5,916,771; and 5,939,598, which are incorporated by reference herein in their entirety. In addition, companies such as Abgenix, Inc. (Freemont, CA) and Genpharm (San Jose, CA) can be engaged to provide human antibodies directed against a selected antigen using technology similar to that described above.

**[0201]** Completely human antibodies which recognize a selected epitope can be generated using a technique referred to as "guided selection." In this approach a selected non-human monoclonal antibody, e.g., a mouse antibody, is used to guide the selection of a completely human antibody recognizing the same epitope. (Jespers et al., *Bio/technology* 12:899-903 (1988)).

**[0202]** Further, antibodies to the polypeptides of the invention can, in turn, be utilized to generate anti-idiotypic antibodies that "mimic" polypeptides of the invention using techniques well known to those skilled in the art. (See, e.g., Greenspan & Bona, *FASEB J.* 7(5):437-444; (1989) and Nissinoff, *J. Immunol.* 147(8):2429-2438 (1991)). For example, antibodies which bind to and competitively inhibit polypeptide multimerization and/or binding of a polypeptide of the invention to a ligand can be used to generate anti-idiotypes that "mimic" the polypeptide multimerization and/or binding domain and, as a consequence, bind to and neutralize polypeptide and/or its ligand. Such neutralizing anti-idiotypes or Fab fragments of such anti-idiotypes can be used in therapeutic regimens to neutralize polypeptide ligand. For example, such anti-idiotypic antibodies can be used to bind a polypeptide of the invention and/or to bind its ligands/receptors, and thereby block its biological activity.

**[0203]     *Polynucleotides Encoding Antibodies***

**[0204]** The invention further provides polynucleotides comprising a nucleotide sequence encoding an antibody of the invention and fragments thereof. The invention also encompasses polynucleotides that hybridize under stringent or lower stringency hybridization conditions, e.g., as defined supra, to polynucleotides that encode an antibody, preferably, that specifically binds to a polypeptide of the invention, preferably, an antibody that binds to a polypeptide having an amino acid sequence in Table 1.

**[0205]** The polynucleotides may be obtained, and the nucleotide sequence of the polynucleotides determined, by any method known in the art. For example, if the nucleotide sequence of the antibody is known, a polynucleotide encoding the antibody may be assembled from chemically synthesized oligonucleotides (e.g., as described in Kutmeier et al., *BioTechniques* 17:242 (1994)), which, briefly, involves the synthesis of overlapping oligonucleotides containing portions of the sequence encoding the antibody, annealing and ligating of those oligonucleotides, and then amplification of the ligated oligonucleotides by PCR.

**[0206]** Alternatively, a polynucleotide encoding an antibody may be generated from nucleic acid from a suitable source. If a clone containing a nucleic acid encoding a particular antibody is not available, but the sequence of the antibody molecule is known, a nucleic acid encoding the immunoglobulin may be chemically synthesized or obtained from a suitable source (e.g., an antibody cDNA library, or a cDNA library generated from, or nucleic acid, preferably poly A+ RNA, isolated from, any tissue or cells expressing the antibody, such as hybridoma cells selected to express an antibody of the invention) by PCR amplification using synthetic primers hybridizable to the 3' and 5' ends of the sequence or by cloning using an oligonucleotide probe specific for the particular gene sequence to identify, for example, a cDNA clone from a cDNA library that encodes the antibody. Amplified nucleic acids generated by PCR may then be cloned into replicable cloning vectors using any method well known in the art.

**[0207]** Once the nucleotide sequence and corresponding amino acid sequence of the antibody is determined, the nucleotide sequence of the antibody may be manipulated using methods well known in the art for the manipulation of nucleotide sequences, e.g., recombinant DNA techniques, site directed mutagenesis, PCR, etc. (see, for example, the techniques described in Sambrook et al., 1990, *Molecular Cloning, A Laboratory Manual*, 2d Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY and Ausubel et al., eds., 1998, *Current Protocols in Molecular Biology*, John Wiley & Sons, NY, which are both incorporated by

reference herein in their entirety), to generate antibodies having a different amino acid sequence, for example to create amino acid substitutions, deletions, and/or insertions.

**[0208]** In a specific embodiment, the amino acid sequence of the heavy and/or light chain variable domains may be inspected to identify the sequences of the complementarity determining regions (CDRs) by methods that are well known in the art, e.g., by comparison to known amino acid sequences of other heavy and light chain variable regions to determine the regions of sequence hypervariability. Using routine recombinant DNA techniques, one or more of the CDRs may be inserted within framework regions, e.g., into human framework regions to humanize a non-human antibody, as described above. The framework regions may be naturally occurring or consensus framework regions, and preferably human framework regions (see, e.g., Chothia et al., *J. Mol. Biol.* 278: 457-479 (1998) for a listing of human framework regions). Preferably, the polynucleotide generated by the combination of the framework regions and CDRs encodes an antibody that specifically binds a polypeptide of the invention. Preferably, as discussed supra, one or more amino acid substitutions may be made within the framework regions, and, preferably, the amino acid substitutions improve binding of the antibody to its antigen. Additionally, such methods may be used to make amino acid substitutions or deletions of one or more variable region cysteine residues participating in an intrachain disulfide bond to generate antibody molecules lacking one or more intrachain disulfide bonds. Other alterations to the polynucleotide are encompassed by the present invention and within the skill of the art.

**[0209]** In addition, techniques developed for the production of "chimeric antibodies" (Morrison et al., *Proc. Natl. Acad. Sci.* 81:851-855 (1984); Neuberger et al., *Nature* 312:604-608 (1984); Takeda et al., *Nature* 314:452-454 (1985)) by splicing genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. As described supra, a chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human immunoglobulin constant region, e.g., humanized antibodies.

**[0210]** Alternatively, techniques described for the production of single chain antibodies (U.S. Patent No. 4,946,778; Bird, *Science* 242:423-42 (1988); Huston et al., *Proc. Natl. Acad. Sci. USA* 85:5879-5883 (1988); and Ward et al., *Nature* 334:544-54 (1989)) can be adapted to produce single chain antibodies. Single chain antibodies are formed by linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain

polypeptide. Techniques for the assembly of functional Fv fragments in *E. coli* may also be used (Skerra et al., Science 242:1038- 1041 (1988)).

**[0211]     *Methods of Producing Antibodies***

**[0212]**     The antibodies of the invention can be produced by any method known in the art for the synthesis of antibodies, in particular, by chemical synthesis or preferably, by recombinant expression techniques.

**[0213]**     Recombinant expression of an antibody of the invention, or fragment, derivative or analog thereof, (e.g., a heavy or light chain of an antibody of the invention or a single chain antibody of the invention), requires construction of an expression vector containing a polynucleotide that encodes the antibody. Once a polynucleotide encoding an antibody molecule or a heavy or light chain of an antibody, or portion thereof (preferably containing the heavy or light chain variable domain), of the invention has been obtained, the vector for the production of the antibody molecule may be produced by recombinant DNA technology using techniques well known in the art. Thus, methods for preparing a protein by expressing a polynucleotide containing an antibody encoding nucleotide sequence are described herein. Methods which are well known to those skilled in the art can be used to construct expression vectors containing antibody coding sequences and appropriate transcriptional and translational control signals. These methods include, for example, in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. The invention, thus, provides replicable vectors comprising a nucleotide sequence encoding an antibody molecule of the invention, or a heavy or light chain thereof, or a heavy or light chain variable domain, operably linked to a promoter. Such vectors may include the nucleotide sequence encoding the constant region of the antibody molecule (see, e.g., PCT Publication WO 86/05807; PCT Publication WO 89/01036; and U.S. Patent No. 5,122,464) and the variable domain of the antibody may be cloned into such a vector for expression of the entire heavy or light chain.

**[0214]**     The expression vector is transferred to a host cell by conventional techniques and the transfected cells are then cultured by conventional techniques to produce an antibody of the invention. Thus, the invention includes host cells containing a polynucleotide encoding an antibody of the invention, or a heavy or light chain thereof, or a single chain antibody of the invention, operably linked to a heterologous promoter. In preferred embodiments for the expression of double-chained antibodies, vectors encoding both the heavy and light chains



may be co-expressed in the host cell for expression of the entire immunoglobulin molecule, as detailed below.

**[0215]** A variety of host-expression vector systems may be utilized to express the antibody molecules of the invention. Such host-expression systems represent vehicles by which the coding sequences of interest may be produced and subsequently purified, but also represent cells which may, when transformed or transfected with the appropriate nucleotide coding sequences, express an antibody molecule of the invention in situ. These include but are not limited to microorganisms such as bacteria (e.g., *E. coli*, *B. subtilis*) transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing antibody coding sequences; yeast (e.g., *Saccharomyces*, *Pichia*) transformed with recombinant yeast expression vectors containing antibody coding sequences; insect cell systems infected with recombinant virus expression vectors (e.g., baculovirus) containing antibody coding sequences; plant cell systems infected with recombinant virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (e.g., Ti plasmid) containing antibody coding sequences; or mammalian cell systems (e.g., COS, CHO, BHK, 293, 3T3 cells) harboring recombinant expression constructs containing promoters derived from the genome of mammalian cells (e.g., metallothionein promoter) or from mammalian viruses (e.g., the adenovirus late promoter; the vaccinia virus 7.5K promoter). Preferably, bacterial cells such as *Escherichia coli*, and more preferably, eukaryotic cells, especially for the expression of whole recombinant antibody molecule, are used for the expression of a recombinant antibody molecule. For example, mammalian cells such as Chinese hamster ovary cells (CHO), in conjunction with a vector such as the major intermediate early gene promoter element from human cytomegalovirus is an effective expression system for antibodies (Foecking et al., *Gene* 45:101 (1986); Cockett et al., *Bio/Technology* 8:2 (1990)).

**[0216]** In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the antibody molecule being expressed. For example, when a large quantity of such a protein is to be produced, for the generation of pharmaceutical compositions of an antibody molecule, vectors which direct the expression of high levels of fusion protein products that are readily purified may be desirable. Such vectors include, but are not limited, to the *E. coli* expression vector pUR278 (Ruther et al., *EMBO J.* 2:1791 (1983)), in which the antibody coding sequence may be ligated individually into the vector in frame with the lac Z coding region so that a fusion protein is produced; pIN vectors

(Inouye & Inouye, *Nucleic Acids Res.* 13:3101-3109 (1985); Van Heeke & Schuster, *J. Biol. Chem.* 24:5503-5509 (1989)); and the like. pGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption and binding to matrix glutathione-agarose beads followed by elution in the presence of free glutathione. The pGEX vectors are designed to include thrombin or factor Xa protease cleavage sites so that the cloned target gene product can be released from the GST moiety.

**[0217]** In an insect system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes. The virus grows in *Spodoptera frugiperda* cells. The antibody coding sequence may be cloned individually into non-essential regions (for example the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (for example the polyhedrin promoter).

**[0218]** In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, the antibody coding sequence of interest may be ligated to an adenovirus transcription/translation control complex, e.g., the late promoter and tripartite leader sequence. This chimeric gene may then be inserted in the adenovirus genome by in vitro or in vivo recombination. Insertion in a non-essential region of the viral genome (e.g., region E1 or E3) will result in a recombinant virus that is viable and capable of expressing the antibody molecule in infected hosts. (e.g., see Logan & Shenk, *Proc. Natl. Acad. Sci. USA* 81:355-359 (1984)). Specific initiation signals may also be required for efficient translation of inserted antibody coding sequences. These signals include the ATG initiation codon and adjacent sequences. Furthermore, the initiation codon must be in phase with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements, transcription terminators, etc. (see Bittner et al., *Methods in Enzymol.* 153:51-544 (1987)).

**[0219]** In addition, a host cell strain may be chosen which modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may be important for the function of the protein. Different host cells have characteristic and specific mechanisms for the post-translational processing and modification of proteins and gene products. Appropriate cell lines or host systems can be chosen to ensure the correct

modification and processing of the foreign protein expressed. To this end, eukaryotic host cells which possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product may be used. Such mammalian host cells include but are not limited to CHO, VERO, BHK, HeLa, COS, MDCK, 293, 3T3, WI38, and in particular, breast cancer cell lines such as, for example, BT483, Hs578T, HTB2, BT20 and T47D, and normal mammary gland cell line such as, for example, CRL7030 and Hs578Bst.

**[0220]** For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines which stably express the antibody molecule may be engineered. Rather than using expression vectors which contain viral origins of replication, host cells can be transformed with DNA controlled by appropriate expression control elements (e.g., promoter, enhancer, sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. Following the introduction of the foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines. This method may advantageously be used to engineer cell lines which express the antibody molecule. Such engineered cell lines may be particularly useful in screening and evaluation of compounds that interact directly or indirectly with the antibody molecule.

**[0221]** A number of selection systems may be used, including but not limited to the herpes simplex virus thymidine kinase (Wigler et al., Cell 11:223 (1977)), hypoxanthine-guanine phosphoribosyltransferase (Szybalska & Szybalski, Proc. Natl. Acad. Sci. USA 48:202 (1992)), and adenine phosphoribosyltransferase (Lowy et al., Cell 22:817 (1980)) genes can be employed in tk-, hgp<sup>r</sup>t- or ap<sup>r</sup>t- cells, respectively. Also, antimetabolite resistance can be used as the basis of selection for the following genes: dhfr, which confers resistance to methotrexate (Wigler et al., Natl. Acad. Sci. USA 77:357 (1980); O'Hare et al., Proc. Natl. Acad. Sci. USA 78:1527 (1981)); gpt, which confers resistance to mycophenolic acid (Mulligan & Berg, Proc. Natl. Acad. Sci. USA 78:2072 (1981)); neo, which confers resistance to the aminoglycoside G-418 Clinical Pharmacy 12:488-505; Wu and Wu, Biotherapy 3:87-95 (1991); Tolstoshev, Ann. Rev. Pharmacol. Toxicol. 32:573-596 (1993); Mulligan, Science 260:926-932 (1993); and Morgan and Anderson, Ann. Rev. Biochem. 62:191-217 (1993); May, 1993, TIB TECH 11(5):155-215); and hyg<sup>r</sup>, which confers resistance to hygromycin

(Santerre et al., *Gene* 30:147 (1984)). Methods commonly known in the art of recombinant DNA technology may be routinely applied to select the desired recombinant clone, and such methods are described, for example, in Ausubel et al. (eds.), *Current Protocols in Molecular Biology*, John Wiley & Sons, NY (1993); Kriegler, *Gene Transfer and Expression, A Laboratory Manual*, Stockton Press, NY (1990); and in Chapters 12 and 13, Dracopoli et al. (eds), *Current Protocols in Human Genetics*, John Wiley & Sons, NY (1994); Colberre-Garapin et al., *J. Mol. Biol.* 150:1 (1981), which are incorporated by reference herein in their entireties.

[0222] The expression levels of an antibody molecule can be increased by vector amplification (for a review, see Bebbington and Hentschel, *The use of vectors based on gene amplification for the expression of cloned genes in mammalian cells in DNA cloning*, Vol.3. (Academic Press, New York, 1987)). When a marker in the vector system expressing antibody is amplifiable, increase in the level of inhibitor present in culture of host cell will increase the number of copies of the marker gene. Since the amplified region is associated with the antibody gene, production of the antibody will also increase (Crouse et al., *Mol. Cell. Biol.* 3:257 (1983)).

[0223] The host cell may be co-transfected with two expression vectors of the invention, the first vector encoding a heavy chain derived polypeptide and the second vector encoding a light chain derived polypeptide. The two vectors may contain identical selectable markers which enable equal expression of heavy and light chain polypeptides. Alternatively, a single vector may be used which encodes, and is capable of expressing, both heavy and light chain polypeptides. In such situations, the light chain should be placed before the heavy chain to avoid an excess of toxic free heavy chain (Proudfoot, *Nature* 322:52 (1986); Kohler, *Proc. Natl. Acad. Sci. USA* 77:2197 (1980)). The coding sequences for the heavy and light chains may comprise cDNA or genomic DNA.

[0224] Once an antibody molecule of the invention has been produced by an animal, chemically synthesized, or recombinantly expressed, it may be purified by any method known in the art for purification of an immunoglobulin molecule, for example, by chromatography (e.g., ion exchange, affinity, particularly by affinity for the specific antigen after Protein A, and sizing column chromatography), centrifugation, differential solubility, or by any other standard technique for the purification of proteins. In addition, the antibodies of the present invention or fragments thereof can be fused to heterologous polypeptide sequences described herein or otherwise known in the art, to facilitate purification.

**[0225]** The present invention encompasses antibodies recombinantly fused or chemically conjugated (including both covalently and non-covalently conjugations) to a polypeptide (or portion thereof, preferably at least 10, 20, 30, 40, 50, 60, 70, 80, 90 or 100 amino acids of the polypeptide) of the present invention to generate fusion proteins. The fusion does not necessarily need to be direct, but may occur through linker sequences. The antibodies may be specific for antigens other than polypeptides (or portion thereof, preferably at least 10, 20, 30, 40, 50, 60, 70, 80, 90 or 100 amino acids of the polypeptide) of the present invention. For example, antibodies may be used to target the polypeptides of the present invention to particular cell types, either in vitro or in vivo, by fusing or conjugating the polypeptides of the present invention to antibodies specific for particular cell surface receptors. Antibodies fused or conjugated to the polypeptides of the present invention may also be used in in vitro immunoassays and purification methods using methods known in the art. See e.g., Harbor et al., *supra*, and PCT publication WO 93/21232; EP 439,095; Naramura et al., *Immunol. Lett.* 39:91-99 (1994); U.S. Patent 5,474,981; Gillies et al., *PNAS* 89:1428-1432 (1992); Fell et al., *J. Immunol.* 146:2446-2452 (1991), which are incorporated by reference in their entireties.

**[0226]** The present invention further includes compositions comprising the polypeptides of the present invention fused or conjugated to antibody domains other than the variable regions. For example, the polypeptides of the present invention may be fused or conjugated to an antibody Fc region, or portion thereof. The antibody portion fused to a polypeptide of the present invention may comprise the constant region, hinge region, CH1 domain, CH2 domain, and CH3 domain or any combination of whole domains or portions thereof. The polypeptides may also be fused or conjugated to the above antibody portions to form multimers. For example, Fc portions fused to the polypeptides of the present invention can form dimers through disulfide bonding between the Fc portions. Higher multimeric forms can be made by fusing the polypeptides to portions of IgA and IgM. Methods for fusing or conjugating the polypeptides of the present invention to antibody portions are known in the art. See, e.g., U.S. Patent Nos. 5,336,603; 5,622,929; 5,359,046; 5,349,053; 5,447,851; 5,112,946; EP 307,434; EP 367,166; PCT publications WO 96/04388; WO 91/06570; Ashkenazi et al., *Proc. Natl. Acad. Sci. USA* 88:10535-10539 (1991); Zheng et al., *J. Immunol.* 154:5590-5600 (1995); and Vil et al., *Proc. Natl. Acad. Sci. USA* 89:11337-11341 (1992) (said references incorporated by reference in their entireties).

**[0227]** As discussed, *supra*, the polypeptides corresponding to a polypeptide, polypeptide fragment, or a variant of a polypeptide sequence shown in Table 1 may be fused or conjugated

to the above antibody portions to increase the in vivo half life of the polypeptides or for use in immunoassays using methods known in the art. Further, any one of the polypeptides shown in Table 1 may be fused or conjugated to the above antibody portions to facilitate purification. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP 394,827; Traunecker et al., *Nature* 331:84-86 (1988)). The polypeptides of the present invention fused or conjugated to an antibody having disulfide-linked dimeric structures (due to the IgG) may also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., *J. Biochem.* 270:3958-3964 (1995)). In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP A 232,262). Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified, would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See, Bennett et al., *J. Molecular Recognition* 8:52-58 (1995); Johanson et al., *J. Biol. Chem.* 270:9459-9471 (1995)).

**[0228]** Moreover, the antibodies or fragments thereof of the present invention can be fused to marker sequences, such as a peptide to facilitate purification. In preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., *Proc. Natl. Acad. Sci. USA* 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Other peptide tags useful for purification include, but are not limited to, the "HA" tag, which corresponds to an epitope derived from the influenza hemagglutinin protein (Wilson et al., *Cell* 37:767 (1984)) and the "flag" tag.

**[0229]** The present invention further encompasses antibodies or fragments thereof conjugated to a diagnostic or therapeutic agent. The antibodies can be used diagnostically to, for example, monitor the development or progression of a tumor as part of a clinical testing procedure to, e.g., determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent

materials, bioluminescent materials, radioactive materials, positron emitting metals using various positron emission tomographies, and nonradioactive paramagnetic metal ions. The detectable substance may be coupled or conjugated either directly to the antibody (or fragment thereof) or indirectly, through an intermediate (such as, for example, a linker known in the art) using techniques known in the art. See, for example, U.S. Patent No. 4,741,900 for metal ions which can be conjugated to antibodies for use as diagnostics according to the present invention. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, beta-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin; and examples of suitable radioactive material include  $^{125}\text{I}$ ,  $^{131}\text{I}$ ,  $^{111}\text{In}$  or  $^{99}\text{Tc}$ .

**[0230]** Further, an antibody or fragment thereof may be conjugated to a therapeutic moiety such as a cytotoxin, e.g., a cytostatic or cytotoxic agent, a therapeutic agent or a radioactive metal ion, e.g., alpha-emitters such as, for example,  $^{213}\text{Bi}$ . A cytotoxin or cytotoxic agent includes any agent that is detrimental to cells. Examples include paclitaxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs thereof. Therapeutic agents include, but are not limited to, antimetabolites (e.g., methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (e.g., mechlorethamine, thioepa chlorambucil, melphalan, carmustine (BSNU) and lomustine (CCNU), cyclophosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (e.g., daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (e.g., dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (e.g., vincristine and vinblastine).

**[0231]** The conjugates of the invention can be used for modifying a given biological response, the therapeutic agent or drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such

as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a protein such as tumor necrosis factor,  $\alpha$ -interferon,  $\beta$ -interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator, an apoptotic agent, e.g., TNF- $\alpha$ , TNF- $\beta$ , AIM I (See, International Publication No. WO 97/33899), AIM II (See, International Publication No. WO 97/34911), Fas Ligand (Takahashi *et al.*, *Int. Immunol.*, 6:1567-1574 (1994)), VEGI (See, International Publication No. WO 99/23105), a thrombotic agent or an anti-angiogenic agent, e.g., angiostatin or endostatin; or, biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophage colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors.

[0232] Antibodies may also be attached to solid supports, which are particularly useful for immunoassays or purification of the target antigen. Such solid supports include, but are not limited to, glass, cellulose, polyacrylamide, nylon, polystyrene, polyvinyl chloride or polypropylene.

[0233] Techniques for conjugating such therapeutic moiety to antibodies are well known, see, e.g., Arnon *et al.*, "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy", in *Monoclonal Antibodies And Cancer Therapy*, Reisfeld *et al.* (eds.), pp. 243-56 (Alan R. Liss, Inc. 1985); Hellstrom *et al.*, "Antibodies For Drug Delivery", in *Controlled Drug Delivery* (2nd Ed.), Robinson *et al.* (eds.), pp. 623-53 (Marcel Dekker, Inc. 1987); Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review", in *Monoclonal Antibodies '84: Biological And Clinical Applications*, Pinchera *et al.* (eds.), pp. 475-506 (1985); "Analysis, Results, And Future Prospective Of The Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", in *Monoclonal Antibodies For Cancer Detection And Therapy*, Baldwin *et al.* (eds.), pp. 303-16 (Academic Press 1985), and Thorpe *et al.*, "The Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates", *Immunol. Rev.* 62:119-58 (1982).

[0234] Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980, which is incorporated herein by reference in its entirety.

[0235] An antibody, with or without a therapeutic moiety conjugated to it, administered alone or in combination with cytotoxic factor(s) and/or cytokine(s) can be used as a therapeutic.



**[0236]     *Immunophenotyping***

**[0237]**     The antibodies of the invention may be utilized for immunophenotyping of cell lines and biological samples. The translation product of the gene of the present invention may be useful as a cell specific marker, or more specifically as a cellular marker that is differentially expressed at various stages of differentiation and/or maturation of particular cell types. Monoclonal antibodies directed against a specific epitope, or combination of epitopes, will allow for the screening of cellular populations expressing the marker. Various techniques can be utilized using monoclonal antibodies to screen for cellular populations expressing the marker(s), and include magnetic separation using antibody-coated magnetic beads, "panning" with antibody attached to a solid matrix (i.e., plate), and flow cytometry (See, e.g., U.S. Patent 5,985,660; and Morrison *et al.*, *Cell*, 96:737-49 (1999)).

**[0238]**     These techniques allow for the screening of particular populations of cells, such as might be found with hematological malignancies (i.e. minimal residual disease (MRD) in acute leukemic patients) and "non-self" cells in transplantations to prevent Graft-versus-Host Disease (GVHD). Alternatively, these techniques allow for the screening of hematopoietic stem and progenitor cells capable of undergoing proliferation and/or differentiation, as might be found in human umbilical cord blood.

**[0239]     *Assays For Antibody Binding***

**[0240]**     The antibodies of the invention may be assayed for immunospecific binding by any method known in the art. The immunoassays which can be used include but are not limited to competitive and non-competitive assay systems using techniques such as western blots, radioimmunoassays, ELISA (enzyme linked immunosorbent assay), "sandwich" immunoassays, immunoprecipitation assays, precipitin reactions, gel diffusion precipitin reactions, immunodiffusion assays, agglutination assays, complement-fixation assays, immunoradiometric assays, fluorescent immunoassays, protein A immunoassays, to name but a few. Such assays are routine and well known in the art (see, e.g., Ausubel et al, eds, 1994, Current Protocols in Molecular Biology, Vol. 1, John Wiley & Sons, Inc., New York, which is incorporated by reference herein in its entirety). Exemplary immunoassays are described briefly below (but are not intended by way of limitation).

**[0241]**     Immunoprecipitation protocols generally comprise lysing a population of cells in a lysis buffer such as RIPA buffer (1% NP-40 or Triton X- 100, 1% sodium deoxycholate, 0.1% SDS, 0.15 M NaCl, 0.01 M sodium phosphate at pH 7.2, 1% Trasylol) supplemented with

protein phosphatase and/or protease inhibitors (e.g., EDTA, PMSF, aprotinin, sodium vanadate), adding the antibody of interest to the cell lysate, incubating for a period of time (e.g., 1-4 hours) at 4° C, adding protein A and/or protein G sepharose beads to the cell lysate, incubating for about an hour or more at 4° C, washing the beads in lysis buffer and resuspending the beads in SDS/sample buffer. The ability of the antibody of interest to immunoprecipitate a particular antigen can be assessed by, e.g., western blot analysis. One of skill in the art would be knowledgeable as to the parameters that can be modified to increase the binding of the antibody to an antigen and decrease the background (e.g., pre-clearing the cell lysate with sepharose beads). For further discussion regarding immunoprecipitation protocols see, e.g., Ausubel et al, eds, 1994, *Current Protocols in Molecular Biology*, Vol. 1, John Wiley & Sons, Inc., New York at 10.16.1.

**[0242]** Western blot analysis generally comprises preparing protein samples, electrophoresis of the protein samples in a polyacrylamide gel (e.g., 8%- 20% SDS-PAGE depending on the molecular weight of the antigen), transferring the protein sample from the polyacrylamide gel to a membrane such as nitrocellulose, PVDF or nylon, blocking the membrane in blocking solution (e.g., PBS with 3% BSA or non-fat milk), washing the membrane in washing buffer (e.g., PBS-Tween 20), blocking the membrane with primary antibody (the antibody of interest) diluted in blocking buffer, washing the membrane in washing buffer, blocking the membrane with a secondary antibody (which recognizes the primary antibody, e.g., an anti-human antibody) conjugated to an enzymatic substrate (e.g., horseradish peroxidase or alkaline phosphatase) or radioactive molecule (e.g., <sup>32</sup>P or <sup>125</sup>I) diluted in blocking buffer, washing the membrane in wash buffer, and detecting the presence of the antigen. One of skill in the art would be knowledgeable as to the parameters that can be modified to increase the signal detected and to reduce the background noise. For further discussion regarding western blot protocols see, e.g., Ausubel et al, eds, 1994, *Current Protocols in Molecular Biology*, Vol. 1, John Wiley & Sons, Inc., New York at 10.8.1.

**[0243]** ELISAs comprise preparing antigen, coating the well of a 96 well microtiter plate with the antigen, adding the antibody of interest conjugated to a detectable compound such as an enzymatic substrate (e.g., horseradish peroxidase or alkaline phosphatase) to the well and incubating for a period of time, and detecting the presence of the antigen. In ELISAs the antibody of interest does not have to be conjugated to a detectable compound; instead, a second antibody (which recognizes the antibody of interest) conjugated to a detectable compound may be added to the well. Further, instead of coating the well with the antigen, the

antibody may be coated to the well. In this case, a second antibody conjugated to a detectable compound may be added following the addition of the antigen of interest to the coated well. One of skill in the art would be knowledgeable as to the parameters that can be modified to increase the signal detected as well as other variations of ELISAs known in the art. For further discussion regarding ELISAs see, e.g., Ausubel et al, eds, 1994, Current Protocols in Molecular Biology, Vol. 1, John Wiley & Sons, Inc., New York at 11.2.1.

**[0244]** The binding affinity of an antibody to an antigen and the off-rate of an antibody-antigen interaction can be determined by competitive binding assays. One example of a competitive binding assay is a radioimmunoassay comprising the incubation of labeled antigen (e.g.,  $^3\text{H}$  or  $^{125}\text{I}$ ) with the antibody of interest in the presence of increasing amounts of unlabeled antigen, and the detection of the antibody bound to the labeled antigen. The affinity of the antibody of interest for a particular antigen and the binding off-rates can be determined from the data by scatchard plot analysis. Competition with a second antibody can also be determined using radioimmunoassays. In this case, the antigen is incubated with antibody of interest conjugated to a labeled compound (e.g.,  $^3\text{H}$  or  $^{125}\text{I}$ ) in the presence of increasing amounts of an unlabeled second antibody.

**[0245]** *Therapeutic Uses*

**[0246]** The present invention is further directed to antibody-based therapies which involve administering antibodies of the invention to an animal, preferably a mammal, and most preferably a human, patient for treating one or more of the disclosed diseases, disorders, or conditions. Therapeutic compounds of the invention include, but are not limited to, antibodies of the invention (including fragments, analogs and derivatives thereof as described herein) and nucleic acids encoding antibodies of the invention (including fragments, analogs and derivatives thereof and anti-idiotypic antibodies as described herein). The antibodies of the invention can be used to treat, inhibit or prevent diseases, disorders or conditions associated with aberrant expression and/or activity of a polypeptide of the invention, including, but not limited to, any one or more of the diseases, disorders, or conditions described herein. The treatment and/or prevention of diseases, disorders, or conditions associated with aberrant expression and/or activity of a polypeptide of the invention includes, but is not limited to, alleviating symptoms associated with those diseases, disorders or conditions. Antibodies of the invention may be provided in pharmaceutically acceptable compositions as known in the art or as described herein.

[0247] A summary of the ways in which the antibodies of the present invention may be used therapeutically includes binding polynucleotides or polypeptides of the present invention locally or systemically in the body or by direct cytotoxicity of the antibody, e.g. as mediated by complement (CDC) or by effector cells (ADCC). Some of these approaches are described in more detail below. Armed with the teachings provided herein, one of ordinary skill in the art will know how to use the antibodies of the present invention for diagnostic, monitoring or therapeutic purposes without undue experimentation.

[0248] The antibodies of this invention may be advantageously utilized in combination with other monoclonal or chimeric antibodies, or with lymphokines or hematopoietic growth factors (such as, e.g., IL-2, IL-3 and IL-7), for example, which serve to increase the number or activity of effector cells which interact with the antibodies.

[0249] The antibodies of the invention may be administered alone or in combination with other types of treatments (e.g., radiation therapy, chemotherapy, hormonal therapy, immunotherapy and anti-tumor agents). Generally, administration of products of a species origin or species reactivity (in the case of antibodies) that is the same species as that of the patient is preferred. Thus, in a preferred embodiment, human antibodies, fragments derivatives, analogs, or nucleic acids, are administered to a human patient for therapy or prophylaxis.

[0250] It is preferred to use high affinity and/or potent in vivo inhibiting and/or neutralizing antibodies against polypeptides or polynucleotides of the present invention, fragments or regions thereof, for both immunoassays directed to and therapy of disorders related to polynucleotides or polypeptides, including fragments thereof, of the present invention. Such antibodies, fragments, or regions, will preferably have an affinity for polynucleotides or polypeptides of the invention, including fragments thereof. Preferred binding affinities include those with a dissociation constant or  $K_d$  less than  $5 \times 10^{-2}$  M,  $10^{-2}$  M,  $5 \times 10^{-3}$  M,  $10^{-3}$  M,  $5 \times 10^{-4}$  M,  $10^{-4}$  M,  $5 \times 10^{-5}$  M,  $10^{-5}$  M,  $5 \times 10^{-6}$  M,  $10^{-6}$  M,  $5 \times 10^{-7}$  M,  $10^{-7}$  M,  $5 \times 10^{-8}$  M,  $10^{-8}$  M,  $5 \times 10^{-9}$  M,  $10^{-9}$  M,  $5 \times 10^{-10}$  M,  $10^{-10}$  M,  $5 \times 10^{-11}$  M,  $10^{-11}$  M,  $5 \times 10^{-12}$  M,  $10^{-12}$  M,  $5 \times 10^{-13}$  M,  $10^{-13}$  M,  $5 \times 10^{-14}$  M,  $10^{-14}$  M,  $5 \times 10^{-15}$  M, and  $10^{-15}$  M.

[0251] *Gene Therapy*

[0252] In a specific embodiment, nucleic acids comprising sequences encoding antibodies or functional derivatives thereof, are administered to treat, inhibit or prevent a disease or disorder associated with aberrant expression and/or activity of a polypeptide of the invention,

by way of gene therapy. Gene therapy refers to therapy performed by the administration to a subject of an expressed or expressible nucleic acid. In this embodiment of the invention, the nucleic acids produce their encoded protein that mediates a therapeutic effect.

**[0253]** Any of the methods for gene therapy available in the art can be used according to the present invention. Exemplary methods are described below.

**[0254]** For general reviews of the methods of gene therapy, see Goldspiel et al., *Clinical Pharmacy* 12:488-505 (1993); Wu and Wu, *Biotherapy* 3:87-95 (1991); Tolstoshev, *Ann. Rev. Pharmacol. Toxicol.* 32:573-596 (1993); Mulligan, *Science* 260:926-932 (1993); and Morgan and Anderson, *Ann. Rev. Biochem.* 62:191-217 (1993); May, *TIBTECH* 11(5):155-215 (1993). Methods commonly known in the art of recombinant DNA technology which can be used are described in Ausubel et al. (eds.), *Current Protocols in Molecular Biology*, John Wiley & Sons, NY (1993); and Kriegler, *Gene Transfer and Expression, A Laboratory Manual*, Stockton Press, NY (1990).

**[0255]** In a preferred aspect, the compound comprises nucleic acid sequences encoding an antibody, said nucleic acid sequences being part of expression vectors that express the antibody or fragments or chimeric proteins or heavy or light chains thereof in a suitable host. In particular, such nucleic acid sequences have promoters operably linked to the antibody coding region, said promoter being inducible or constitutive, and, optionally, tissue-specific. In another particular embodiment, nucleic acid molecules are used in which the antibody coding sequences and any other desired sequences are flanked by regions that promote homologous recombination at a desired site in the genome, thus providing for intrachromosomal expression of the antibody encoding nucleic acids (Koller and Smithies, *Proc. Natl. Acad. Sci. USA* 86:8932-8935 (1989); Zijlstra et al., *Nature* 342:435-438 (1989). In specific embodiments, the expressed antibody molecule is a single chain antibody; alternatively, the nucleic acid sequences include sequences encoding both the heavy and light chains, or fragments thereof, of the antibody.

**[0256]** Delivery of the nucleic acids into a patient may be either direct, in which case the patient is directly exposed to the nucleic acid or nucleic acid-carrying vectors, or indirect, in which case, cells are first transformed with the nucleic acids in vitro, then transplanted into the patient. These two approaches are known, respectively, as in vivo or ex vivo gene therapy.

**[0257]** In a specific embodiment, the nucleic acid sequences are directly administered in vivo, where it is expressed to produce the encoded product. This can be accomplished by any of numerous methods known in the art, e.g., by constructing them as part of an appropriate

nucleic acid expression vector and administering it so that they become intracellular, e.g., by infection using defective or attenuated retrovirals or other viral vectors (see U.S. Patent No. 4,980,286), or by direct injection of naked DNA, or by use of microparticle bombardment (e.g., a gene gun; Biolistic, Dupont), or coating with lipids or cell-surface receptors or transfecting agents, encapsulation in liposomes, microparticles, or microcapsules, or by administering them in linkage to a peptide which is known to enter the nucleus, by administering it in linkage to a ligand subject to receptor-mediated endocytosis (see, e.g., Wu and Wu, *J. Biol. Chem.* 262:4429-4432 (1987)) (which can be used to target cell types specifically expressing the receptors), etc. In another embodiment, nucleic acid-ligand complexes can be formed in which the ligand comprises a fusogenic viral peptide to disrupt endosomes, allowing the nucleic acid to avoid lysosomal degradation. In yet another embodiment, the nucleic acid can be targeted in vivo for cell specific uptake and expression, by targeting a specific receptor (see, e.g., PCT Publications WO 92/06180; WO 92/22635; WO92/20316; WO93/14188, WO 93/20221). Alternatively, the nucleic acid can be introduced intracellularly and incorporated within host cell DNA for expression, by homologous recombination (Koller and Smithies, *Proc. Natl. Acad. Sci. USA* 86:8932-8935 (1989); Zijlstra et al., *Nature* 342:435-438 (1989)).

**[0258]** In a specific embodiment, viral vectors that contains nucleic acid sequences encoding an antibody of the invention are used. For example, a retroviral vector can be used (see Miller et al., *Meth. Enzymol.* 217:581-599 (1993)). These retroviral vectors contain the components necessary for the correct packaging of the viral genome and integration into the host cell DNA. The nucleic acid sequences encoding the antibody to be used in gene therapy are cloned into one or more vectors, which facilitates delivery of the gene into a patient. More detail about retroviral vectors can be found in Boesen et al., *Biotherapy* 6:291-302 (1994), which describes the use of a retroviral vector to deliver the *mdr1* gene to hematopoietic stem cells in order to make the stem cells more resistant to chemotherapy. Other references illustrating the use of retroviral vectors in gene therapy are: Clowes et al., *J. Clin. Invest.* 93:644-651 (1994); Kiem et al., *Blood* 83:1467-1473 (1994); Salmons and Gunzberg, *Human Gene Therapy* 4:129-141 (1993); and Grossman and Wilson, *Curr. Opin. in Genetics and Devel.* 3:110-114 (1993).

**[0259]** Adenoviruses are other viral vectors that can be used in gene therapy. Adenoviruses are especially attractive vehicles for delivering genes to respiratory epithelia. Adenoviruses naturally infect respiratory epithelia where they cause a mild disease. Other targets for adenovirus-based delivery systems are liver, the central nervous system, endothelial

cells, and muscle. Adenoviruses have the advantage of being capable of infecting non-dividing cells. Kozarsky and Wilson, *Current Opinion in Genetics and Development* 3:499-503 (1993) present a review of adenovirus-based gene therapy. Bout et al., *Human Gene Therapy* 5:3-10 (1994) demonstrated the use of adenovirus vectors to transfer genes to the respiratory epithelia of rhesus monkeys. Other instances of the use of adenoviruses in gene therapy can be found in Rosenfeld et al., *Science* 252:431-434 (1991); Rosenfeld et al., *Cell* 68:143-155 (1992); Mastrangeli et al., *J. Clin. Invest.* 91:225-234 (1993); PCT Publication WO94/12649; and Wang, et al., *Gene Therapy* 2:775-783 (1995). In a preferred embodiment, adenovirus vectors are used.

[0260] Adeno-associated virus (AAV) has also been proposed for use in gene therapy (Walsh et al., *Proc. Soc. Exp. Biol. Med.* 204:289-300 (1993); U.S. Patent No. 5,436,146).

[0261] Another approach to gene therapy involves transferring a gene to cells in tissue culture by such methods as electroporation, lipofection, calcium phosphate mediated transfection, or viral infection. Usually, the method of transfer includes the transfer of a selectable marker to the cells. The cells are then placed under selection to isolate those cells that have taken up and are expressing the transferred gene. Those cells are then delivered to a patient.

[0262] In this embodiment, the nucleic acid is introduced into a cell prior to administration in vivo of the resulting recombinant cell. Such introduction can be carried out by any method known in the art, including but not limited to transfection, electroporation, microinjection, infection with a viral or bacteriophage vector containing the nucleic acid sequences, cell fusion, chromosome-mediated gene transfer, microcell-mediated gene transfer, spheroplast fusion, etc. Numerous techniques are known in the art for the introduction of foreign genes into cells (see, e.g., Loeffler and Behr, *Meth. Enzymol.* 217:599-618 (1993); Cohen et al., *Meth. Enzymol.* 217:618-644 (1993); Cline, *Pharmac. Ther.* 29:69-92m (1985) and may be used in accordance with the present invention, provided that the necessary developmental and physiological functions of the recipient cells are not disrupted. The technique should provide for the stable transfer of the nucleic acid to the cell, so that the nucleic acid is expressible by the cell and preferably heritable and expressible by its cell progeny.

[0263] The resulting recombinant cells can be delivered to a patient by various methods known in the art. Recombinant blood cells (e.g., hematopoietic stem or progenitor cells) are preferably administered intravenously. The amount of cells envisioned for use depends on the desired effect, patient state, etc., and can be determined by one skilled in the art.

**[0264]** Cells into which a nucleic acid can be introduced for purposes of gene therapy encompass any desired, available cell type, and include but are not limited to epithelial cells, endothelial cells, keratinocytes, fibroblasts, muscle cells, hepatocytes; blood cells such as Tlymphocytes, Blymphocytes, monocytes, macrophages, neutrophils, eosinophils, megakaryocytes, granulocytes; various stem or progenitor cells, in particular hematopoietic stem or progenitor cells, e.g., as obtained from bone marrow, umbilical cord blood, peripheral blood, fetal liver, etc.

**[0265]** In a preferred embodiment, the cell used for gene therapy is autologous to the patient.

**[0266]** In an embodiment in which recombinant cells are used in gene therapy, nucleic acid sequences encoding an antibody are introduced into the cells such that they are expressible by the cells or their progeny, and the recombinant cells are then administered in vivo for therapeutic effect. In a specific embodiment, stem or progenitor cells are used. Any stem and/or progenitor cells which can be isolated and maintained in vitro can potentially be used in accordance with this embodiment of the present invention (see e.g. PCT Publication WO 94/08598; Stemple and Anderson, Cell 71:973-985 (1992); Rheinwald, Meth. Cell Bio. 21A:229 (1980); and Pittelkow and Scott, Mayo Clinic Proc. 61:771 (1986)).

**[0267]** In a specific embodiment, the nucleic acid to be introduced for purposes of gene therapy comprises an inducible promoter operably linked to the coding region, such that expression of the nucleic acid is controllable by controlling the presence or absence of the appropriate inducer of transcription. Demonstration of Therapeutic or Prophylactic Activity

**[0268]** The compounds or pharmaceutical compositions of the invention are preferably tested in vitro, and then in vivo for the desired therapeutic or prophylactic activity, prior to use in humans. For example, in vitro assays to demonstrate the therapeutic or prophylactic utility of a compound or pharmaceutical composition include, the effect of a compound on a cell line or a patient tissue sample. The effect of the compound or composition on the cell line and/or tissue sample can be determined utilizing techniques known to those of skill in the art including, but not limited to, rosette formation assays and cell lysis assays. In accordance with the invention, in vitro assays which can be used to determine whether administration of a specific compound is indicated, include in vitro cell culture assays in which a patient tissue sample is grown in culture, and exposed to or otherwise administered a compound, and the effect of such compound upon the tissue sample is observed.



**[0269]     *Therapeutic/Prophylactic Administration and Composition***

**[0270]**     The invention provides methods of treatment, inhibition and prophylaxis by administration to a subject of an effective amount of a compound or pharmaceutical composition of the invention, preferably an antibody of the invention. In a preferred aspect, the compound is substantially purified (e.g., substantially free from substances that limit its effect or produce undesired side-effects). The subject is preferably an animal, including but not limited to animals such as cows, pigs, horses, chickens, cats, dogs, etc., and is preferably a mammal, and most preferably human.

**[0271]**     Formulations and methods of administration that can be employed when the compound comprises a nucleic acid or an immunoglobulin are described above; additional appropriate formulations and routes of administration can be selected from among those described herein below.

**[0272]**     Various delivery systems are known and can be used to administer a compound of the invention, e.g., encapsulation in liposomes, microparticles, microcapsules, recombinant cells capable of expressing the compound, receptor-mediated endocytosis (see, e.g., Wu and Wu, J. Biol. Chem. 262:4429-4432 (1987)), construction of a nucleic acid as part of a retroviral or other vector, etc. Methods of introduction include but are not limited to intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, intranasal, epidural, and oral routes. The compounds or compositions may be administered by any convenient route, for example by infusion or bolus injection, by absorption through epithelial or mucocutaneous linings (e.g., oral mucosa, rectal and intestinal mucosa, etc.) and may be administered together with other biologically active agents. Administration can be systemic or local. In addition, it may be desirable to introduce the pharmaceutical compounds or compositions of the invention into the central nervous system by any suitable route, including intraventricular and intrathecal injection; intraventricular injection may be facilitated by an intraventricular catheter, for example, attached to a reservoir, such as an Ommaya reservoir. Pulmonary administration can also be employed, e.g., by use of an inhaler or nebulizer, and formulation with an aerosolizing agent.

**[0273]**     In a specific embodiment, it may be desirable to administer the pharmaceutical compounds or compositions of the invention locally to the area in need of treatment; this may be achieved by, for example, and not by way of limitation, local infusion during surgery, topical application, e.g., in conjunction with a wound dressing after surgery, by injection, by means of a catheter, by means of a suppository, or by means of an implant, said implant being

of a porous, non-porous, or gelatinous material, including membranes, such as sialastic membranes, or fibers. Preferably, when administering a protein, including an antibody, of the invention, care must be taken to use materials to which the protein does not absorb.

**[0274]** In another embodiment, the compound or composition can be delivered in a vesicle, in particular a liposome (see Langer, *Science* 249:1527-1533 (1990); Treat et al., in *Liposomes in the Therapy of Infectious Disease and Cancer*, Lopez-Berestein and Fidler (eds.), Liss, New York, pp. 353- 365 (1989); Lopez-Berestein, *ibid.*, pp. 317-327; see generally *ibid.*)

**[0275]** In yet another embodiment, the compound or composition can be delivered in a controlled release system. In one embodiment, a pump may be used (see Langer, *supra*; Sefton, *CRC Crit. Ref. Biomed. Eng.* 14:201 (1987); Buchwald et al., *Surgery* 88:507 (1980); Saudek et al., *N. Engl. J. Med.* 321:574 (1989)). In another embodiment, polymeric materials can be used (see *Medical Applications of Controlled Release*, Langer and Wise (eds.), CRC Pres., Boca Raton, Florida (1974); *Controlled Drug Bioavailability, Drug Product Design and Performance*, Smolen and Ball (eds.), Wiley, New York (1984); Ranger and Peppas, J., *Macromol. Sci. Rev. Macromol. Chem.* 23:61 (1983); see also Levy et al., *Science* 228:190 (1985); During et al., *Ann. Neurol.* 25:351 (1989); Howard et al., *J.Neurosurg.* 71:105 (1989)). In yet another embodiment, a controlled release system can be placed in proximity of the therapeutic target, i.e., the brain, thus requiring only a fraction of the systemic dose (see, e.g., Goodson, in *Medical Applications of Controlled Release*, *supra*, vol. 2, pp. 115-138 (1984)).

**[0276]** Other controlled release systems are discussed in the review by Langer (*Science* 249:1527-1533 (1990)).

**[0277]** In a specific embodiment where the compound of the invention is a nucleic acid encoding a protein, the nucleic acid can be administered *in vivo* to promote expression of its encoded protein, by constructing it as part of an appropriate nucleic acid expression vector and administering it so that it becomes intracellular, e.g., by use of a retroviral vector (see U.S. Patent No. 4,980,286), or by direct injection, or by use of microparticle bombardment (e.g., a gene gun; Biolistic, Dupont), or coating with lipids or cell-surface receptors or transfecting agents, or by administering it in linkage to a homeobox- like peptide which is known to enter the nucleus (see e.g., Joliot et al., *Proc. Natl. Acad. Sci. USA* 88:1864-1868 (1991)), etc. Alternatively, a nucleic acid can be introduced intracellularly and incorporated within host cell DNA for expression, by homologous recombination.

**[0278]** The present invention also provides pharmaceutical compositions. Such compositions comprise a therapeutically effective amount of a compound, and a

pharmaceutically acceptable carrier. In a specific embodiment, the term "pharmaceutically acceptable" means approved by a regulatory agency of the Federal or a state government or listed in the U.S. Pharmacopeia or other generally recognized pharmacopeia for use in animals, and more particularly in humans. The term "carrier" refers to a diluent, adjuvant, excipient, or vehicle with which the therapeutic is administered. Such pharmaceutical carriers can be sterile liquids, such as water and oils, including those of petroleum, animal, vegetable or synthetic origin, such as peanut oil, soybean oil, mineral oil, sesame oil and the like. Water is a preferred carrier when the pharmaceutical composition is administered intravenously. Saline solutions and aqueous dextrose and glycerol solutions can also be employed as liquid carriers, particularly for injectable solutions. Suitable pharmaceutical excipients include starch, glucose, lactose, sucrose, gelatin, malt, rice, flour, chalk, silica gel, sodium stearate, glycerol monostearate, talc, sodium chloride, dried skim milk, glycerol, propylene, glycol, water, ethanol and the like. The composition, if desired, can also contain minor amounts of wetting or emulsifying agents, or pH buffering agents. These compositions can take the form of solutions, suspensions, emulsion, tablets, pills, capsules, powders, sustained-release formulations and the like. The composition can be formulated as a suppository, with traditional binders and carriers such as triglycerides. Oral formulation can include standard carriers such as pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, etc. Examples of suitable pharmaceutical carriers are described in "Remington's Pharmaceutical Sciences" by E.W. Martin. Such compositions will contain a therapeutically effective amount of the compound, preferably in purified form, together with a suitable amount of carrier so as to provide the form for proper administration to the patient. The formulation should suit the mode of administration.

[0279] In a preferred embodiment, the composition is formulated in accordance with routine procedures as a pharmaceutical composition adapted for intravenous administration to human beings. Typically, compositions for intravenous administration are solutions in sterile isotonic aqueous buffer. Where necessary, the composition may also include a solubilizing agent and a local anesthetic such as lignocaine to ease pain at the site of the injection. Generally, the ingredients are supplied either separately or mixed together in unit dosage form, for example, as a dry lyophilized powder or water free concentrate in a hermetically sealed container such as an ampoule or sachette indicating the quantity of active agent. Where the composition is to be administered by infusion, it can be dispensed with an infusion bottle containing sterile pharmaceutical grade water or saline. Where the composition is

administered by injection, an ampoule of sterile water for injection or saline can be provided so that the ingredients may be mixed prior to administration.

**[0280]** The compounds of the invention can be formulated as neutral or salt forms. Pharmaceutically acceptable salts include those formed with anions such as those derived from hydrochloric, phosphoric, acetic, oxalic, tartaric acids, etc., and those formed with cations such as those derived from sodium, potassium, ammonium, calcium, ferric hydroxides, isopropylamine, triethylamine, 2-ethylamino ethanol, histidine, procaine, etc.

**[0281]** The amount of the compound of the invention which will be effective in the treatment, inhibition and prevention of a disease or disorder associated with aberrant expression and/or activity of a polypeptide of the invention can be determined by standard clinical techniques. In addition, in vitro assays may optionally be employed to help identify optimal dosage ranges. The precise dose to be employed in the formulation will also depend on the route of administration, and the seriousness of the disease or disorder, and should be decided according to the judgment of the practitioner and each patient's circumstances. Effective doses may be extrapolated from dose-response curves derived from in vitro or animal model test systems.

**[0282]** For antibodies, the dosage administered to a patient is typically 0.1 mg/kg to 100 mg/kg of the patient's body weight. Preferably, the dosage administered to a patient is between 0.1 mg/kg and 20 mg/kg of the patient's body weight, more preferably 1 mg/kg to 10 mg/kg of the patient's body weight. Generally, human antibodies have a longer half-life within the human body than antibodies from other species due to the immune response to the foreign polypeptides. Thus, lower dosages of human antibodies and less frequent administration is often possible. Further, the dosage and frequency of administration of antibodies of the invention may be reduced by enhancing uptake and tissue penetration (e.g., into the brain) of the antibodies by modifications such as, for example, lipidation.

**[0283]** The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Optionally associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. **Diagnosis and Imaging**

**[0284]** Labeled antibodies, and derivatives and analogs thereof, which specifically bind to a polypeptide of interest can be used for diagnostic purposes to detect, diagnose, or monitor

diseases and/or disorders associated with the aberrant expression and/or activity of a polypeptide of the invention. The invention provides for the detection of aberrant expression of a polypeptide of interest, comprising (a) assaying the expression of the polypeptide of interest in cells or body fluid of an individual using one or more antibodies specific to the polypeptide interest and (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of aberrant expression.

**[0285]** The invention provides a diagnostic assay for diagnosing a disorder, comprising (a) assaying the expression of the polypeptide of interest in cells or body fluid of an individual using one or more antibodies specific to the polypeptide interest and (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a particular disorder. With respect to cancer, the presence of a relatively high amount of transcript in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

**[0286]** Antibodies of the invention can be used to assay protein levels in a biological sample using classical immunohistological methods known to those of skill in the art (e.g., see Jalkanen, et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, et al., J. Cell. Biol. 105:3087-3096 (1987)). Other antibody-based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase; radioisotopes, such as iodine (125I, 121I), carbon (14C), sulfur (35S), tritium (3H), indium (112In), and technetium (99Tc); luminescent labels, such as luminol; and fluorescent labels, such as fluorescein and rhodamine, and biotin.

**[0287]** One aspect of the invention is the detection and diagnosis of a disease or disorder associated with aberrant expression of a polypeptide of interest in an animal, preferably a mammal and most preferably a human. In one embodiment, diagnosis comprises: a) administering (for example, parenterally, subcutaneously, or intraperitoneally) to a subject an effective amount of a labeled molecule which specifically binds to the polypeptide of interest; b) waiting for a time interval following the administering for permitting the labeled molecule

to preferentially concentrate at sites in the subject where the polypeptide is expressed (and for unbound labeled molecule to be cleared to background level); c) determining background level; and d) detecting the labeled molecule in the subject, such that detection of labeled molecule above the background level indicates that the subject has a particular disease or disorder associated with aberrant expression of the polypeptide of interest. Background level can be determined by various methods including, comparing the amount of labeled molecule detected to a standard value previously determined for a particular system.

**[0288]** It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of <sup>99m</sup>Tc. The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).

**[0289]** Depending on several variables, including the type of label used and the mode of administration, the time interval following the administration for permitting the labeled molecule to preferentially concentrate at sites in the subject and for unbound labeled molecule to be cleared to background level is 6 to 48 hours or 6 to 24 hours or 6 to 12 hours. In another embodiment the time interval following administration is 5 to 20 days or 5 to 10 days.

**[0290]** In an embodiment, monitoring of the disease or disorder is carried out by repeating the method for diagnosing the disease or disease, for example, one month after initial diagnosis, six months after initial diagnosis, one year after initial diagnosis, etc.

**[0291]** Presence of the labeled molecule can be detected in the patient using methods known in the art for in vivo scanning. These methods depend upon the type of label used. Skilled artisans will be able to determine the appropriate method for detecting a particular label. Methods and devices that may be used in the diagnostic methods of the invention include, but are not limited to, computed tomography (CT), whole body scan such as position emission tomography (PET), magnetic resonance imaging (MRI), and sonography.

**[0292]** In a specific embodiment, the molecule is labeled with a radioisotope and is detected in the patient using a radiation responsive surgical instrument (Thurston et al., U.S. Patent No. 5,441,050). In another embodiment, the molecule is labeled with a fluorescent

compound and is detected in the patient using a fluorescence responsive scanning instrument. In another embodiment, the molecule is labeled with a positron emitting metal and is detected in the patient using positron emission-tomography. In yet another embodiment, the molecule is labeled with a paramagnetic label and is detected in a patient using magnetic resonance imaging (MRI).

**[0293]     *Kits***

**[0294]**     The present invention provides kits that can be used in the above methods. In one embodiment, a kit comprises an antibody of the invention, preferably a purified antibody, in one or more containers. In a specific embodiment, the kits of the present invention contain a substantially isolated polypeptide comprising an epitope which is specifically immunoreactive with an antibody included in the kit. Preferably, the kits of the present invention further comprise a control antibody which does not react with the polypeptide of interest. In another specific embodiment, the kits of the present invention contain a means for detecting the binding of an antibody to a polypeptide of interest (e.g., the antibody may be conjugated to a detectable substrate such as a fluorescent compound, an enzymatic substrate, a radioactive compound or a luminescent compound, or a second antibody which recognizes the first antibody may be conjugated to a detectable substrate).

**[0295]**     In another specific embodiment of the present invention, the kit is a diagnostic kit for use in screening serum containing antibodies specific against proliferative and/or cancerous polynucleotides and polypeptides. Such a kit may include a control antibody that does not react with the polypeptide of interest. Such a kit may include a substantially isolated polypeptide antigen comprising an epitope which is specifically immunoreactive with at least one anti-polypeptide antigen antibody. Further, such a kit includes means for detecting the binding of said antibody to the antigen (e.g., the antibody may be conjugated to a fluorescent compound such as fluorescein or rhodamine which can be detected by flow cytometry). In specific embodiments, the kit may include a recombinantly produced or chemically synthesized polypeptide antigen. The polypeptide antigen of the kit may also be attached to a solid support.

**[0296]**     In a more specific embodiment the detecting means of the above-described kit includes a solid support to which said polypeptide antigen is attached. Such a kit may also include a non-attached reporter-labeled anti-human antibody. In this embodiment, binding of

the antibody to the polypeptide antigen can be detected by binding of the said reporter-labeled antibody.

[0297] In an additional embodiment, the invention includes a diagnostic kit for use in screening serum containing antigens of the polypeptide of the invention. The diagnostic kit includes a substantially isolated antibody specifically immunoreactive with polypeptide or polynucleotide antigens, and means for detecting the binding of the polynucleotide or polypeptide antigen to the antibody. In one embodiment, the antibody is attached to a solid support. In a specific embodiment, the antibody may be a monoclonal antibody. The detecting means of the kit may include a second, labeled monoclonal antibody. Alternatively, or in addition, the detecting means may include a labeled, competing antigen.

[0298] In one diagnostic configuration, test serum is reacted with a solid phase reagent having a surface-bound antigen obtained by the methods of the present invention. After binding with specific antigen antibody to the reagent and removing unbound serum components by washing, the reagent is reacted with reporter-labeled anti-human antibody to bind reporter to the reagent in proportion to the amount of bound anti-antigen antibody on the solid support. The reagent is again washed to remove unbound labeled antibody, and the amount of reporter associated with the reagent is determined. Typically, the reporter is an enzyme which is detected by incubating the solid phase in the presence of a suitable fluorometric, luminescent or colorimetric substrate (Sigma, St. Louis, MO).

[0299] The solid surface reagent in the above assay is prepared by known techniques for attaching protein material to solid support material, such as polymeric beads, dip sticks, 96-well plate or filter material. These attachment methods generally include non-specific adsorption of the protein to the support or covalent attachment of the protein, typically through a free amine group, to a chemically reactive group on the solid support, such as an activated carboxyl, hydroxyl, or aldehyde group. Alternatively, streptavidin coated plates can be used in conjunction with biotinylated antigen(s).

[0300] Thus, the invention provides an assay system or kit for carrying out this diagnostic method. The kit generally includes a support with surface-bound recombinant antigens, and a reporter-labeled anti-human antibody for detecting surface-bound anti-antigen antibody.

[0301] *Diagnostic Assays*

[0302] The present invention further relates to methods for assaying staphylococcal infection in an animal (e.g., a mammal, including but not limited to a human) by detecting the



expression of genes encoding staphylococcal polypeptides of the present invention. The methods comprise analyzing tissue or body fluid from the animal for *Staphylococcus*-specific antibodies, nucleic acids, or proteins. Analysis of nucleic acid specific to *Staphylococcus* is assayed by PCR or hybridization techniques using nucleic acid sequences of the present invention as either hybridization probes or primers. See, e.g., Sambrook et al. Molecular cloning: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 2nd ed., 1989, page 54 reference); Ereemeeva et al. (1994) J. Clin. Microbiol. 32:803-810 (describing differentiation among spotted fever group *Rickettsiae* species by analysis of restriction fragment length polymorphism of PCR-amplified DNA) and Chen et al. 1994 J. Clin. Microbiol. 32:589-595 (detecting bacterial nucleic acids *via* PCR).

[0303] Where diagnosis of a disease state related to infection with *Staphylococcus* has already been made, the present invention is useful for monitoring progression or regression of the disease state by measuring the amount of *Staphylococcus* cells present in a patient or whereby patients exhibiting enhanced *Staphylococcus* gene expression will experience a worse clinical outcome relative to patients expressing these gene(s) at a lower level.

[0304] By "biological sample" is intended any biological sample obtained from an animal, cell line, tissue culture, or other source which contains *Staphylococcus* polypeptide, mRNA, or DNA. Biological samples include body fluids (such as saliva, blood, plasma, urine, mucus, synovial fluid, etc.) tissues (such as muscle, skin, and cartilage) and any other biological source suspected of containing *Staphylococcus* polypeptides or nucleic acids. Methods for obtaining biological samples such as tissue are well known in the art.

[0305] The present invention is useful for detecting diseases related to *Staphylococcus* infections in animals. Preferred animals include monkeys, apes, cats, dogs, birds, cows, pigs, mice, horses, rabbits and humans. Particularly preferred are humans.

[0306] Total RNA can be isolated from a biological sample using any suitable technique such as the single-step guanidinium-thiocyanate-phenol-chloroform method described in Chomczynski et al. (1987) Anal. Biochem. 162:156-159. mRNA encoding *Staphylococcus* polypeptides having sufficient homology to the nucleic acid sequences identified in Table 1 to allow for hybridization between complementary sequences are then assayed using any appropriate method. These include Northern blot analysis, S1 nuclease mapping, the polymerase chain reaction (PCR), reverse transcription in combination with the polymerase chain reaction (RT-PCR), and reverse transcription in combination with the ligase chain reaction (RT-LCR).

[0307] Northern blot analysis can be performed as described in Harada et al. (1990) Cell 63:303-312. Briefly, total RNA is prepared from a biological sample as described above. For the Northern blot, the RNA is denatured in an appropriate buffer (such as glyoxal/dimethyl sulfoxide/sodium phosphate buffer), subjected to agarose gel electrophoresis, and transferred onto a nitrocellulose filter. After the RNAs have been linked to the filter by a UV linker, the filter is prehybridized in a solution containing formamide, SSC, Denhardt's solution, denatured salmon sperm, SDS, and sodium phosphate buffer. A *S. aureus* polynucleotide sequence shown in Table 1 labeled according to any appropriate method (such as the <sup>32</sup>P-multiprimed DNA labeling system (Amersham)) is used as probe. After hybridization overnight, the filter is washed and exposed to x-ray film. DNA for use as probe according to the present invention is described in the sections above and will preferably at least 15 nucleotides in length.

[0308] S1 mapping can be performed as described in Fujita et al. (1987) Cell 49:357-367. To prepare probe DNA for use in S1 mapping, the sense strand of an above-described *S. aureus* DNA sequence of the present invention is used as a template to synthesize labeled antisense DNA. The antisense DNA can then be digested using an appropriate restriction endonuclease to generate further DNA probes of a desired length. Such antisense probes are useful for visualizing protected bands corresponding to the target mRNA (*i.e.*, mRNA encoding polypeptides of the present invention).

[0309] Levels of mRNA encoding *Staphylococcus* polypeptides are assayed, for *e.g.*, using the RT-PCR method described in Makino et al. (1990) Technique 2:295-301. By this method, the radioactivities of the "amplicons" in the polyacrylamide gel bands are linearly related to the initial concentration of the target mRNA. Briefly, this method involves adding total RNA isolated from a biological sample in a reaction mixture containing a RT primer and appropriate buffer. After incubating for primer annealing, the mixture can be supplemented with a RT buffer, dNTPs, DTT, RNase inhibitor and reverse transcriptase. After incubation to achieve reverse transcription of the RNA, the RT products are then subject to PCR using labeled primers. Alternatively, rather than labeling the primers, a labeled dNTP can be included in the PCR reaction mixture. PCR amplification can be performed in a DNA thermal cycler according to conventional techniques. After a suitable number of rounds to achieve amplification, the PCR reaction mixture is electrophoresed on a polyacrylamide gel. After drying the gel, the radioactivity of the appropriate bands (corresponding to the mRNA encoding the *Staphylococcus* polypeptides of the present invention) are quantified using an imaging analyzer. RT and PCR reaction ingredients and conditions, reagent and gel

concentrations, and labeling methods are well known in the art. Variations on the RT-PCR method will be apparent to the skilled artisan. Other PCR methods that can detect the nucleic acid of the present invention can be found in PCR PRIMER: A LABORATORY MANUAL (C.W. Dieffenbach et al. eds., Cold Spring Harbor Lab Press, 1995).

**[0310]** The polynucleotides of the present invention, including both DNA and RNA, may be used to detect polynucleotides of the present invention or *Staphylococcus* species including *S. aureus* using bio chip technology. The present invention includes both high density chip arrays ( $>1000$  oligonucleotides per  $\text{cm}^2$ ) and low density chip arrays ( $<1000$  oligonucleotides per  $\text{cm}^2$ ). Bio chips comprising arrays of polynucleotides of the present invention may be used to detect *Staphylococcus* species, including *S. aureus*, in biological and environmental samples and to diagnose an animal, including humans, with an *S. aureus* or other *Staphylococcus* infection. The bio chips of the present invention may comprise polynucleotide sequences of other pathogens including bacteria, viral, parasitic, and fungal polynucleotide sequences, in addition to the polynucleotide sequences of the present invention, for use in rapid differential pathogenic detection and diagnosis. The bio chips can also be used to monitor an *S. aureus* or other *Staphylococcus* infections and to monitor the genetic changes (deletions, insertions, mismatches, etc.) in response to drug therapy in the clinic and drug development in the laboratory. The bio chip technology comprising arrays of polynucleotides of the present invention may also be used to simultaneously monitor the expression of a multiplicity of genes, including those of the present invention. The polynucleotides used to comprise a selected array may be specified in the same manner as for the fragments, i.e, by their 5' and 3' positions or length in contiguous base pairs and include from. Methods and particular uses of the polynucleotides of the present invention to detect *Staphylococcus* species, including *S. aureus*, using bio chip technology include those known in the art and those of: U.S. Patent Nos. 5510270, 5545531, 5445934, 5677195, 5532128, 5556752, 5527681, 5451683, 5424186, 5607646, 5658732 and World Patent Nos. WO/9710365, WO/9511995, WO/9743447, WO/9535505, each incorporated herein in their entireties.

**[0311]** Biosensors using the polynucleotides of the present invention may also be used to detect, diagnose, and monitor *S. aureus* or other *Staphylococcus* species and infections thereof. Biosensors using the polynucleotides of the present invention may also be used to detect particular polynucleotides of the present invention. Biosensors using the polynucleotides of the present invention may also be used to monitor the genetic changes (deletions, insertions, mismatches, etc.) in response to drug therapy in the clinic and drug development in the

laboratory. Methods and particular uses of the polynucleotides of the present invention to detect *Staphylococcus* species, including *S. aureus*, using biosensors include those known in the art and those of: U.S. Patent Nos 5721102, 5658732, 5631170, and World Patent Nos. WO97/35011, WO/9720203, each incorporated herein in their entireties.

[0312] Thus, the present invention includes both bio chips and biosensors comprising polynucleotides of the present invention and methods of their use.

[0313] A preferred composition of matter comprises isolated nucleic acid molecules wherein the nucleotide sequences of said nucleic acid molecules comprise a bio chip or biosensor of at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 100, 150, 200, 250, 300, 500, 1000, 2000, 3000 or 4000 nucleotide sequences, wherein at least one sequence in said DNA bio chip or biosensor is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a *S. aureus* polynucleotide shown in Table 1. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

[0314] Assaying *Staphylococcus* polypeptide levels in a biological sample can occur using any art-known method, such as antibody-based techniques. For example, *Staphylococcus* polypeptide expression in tissues can be studied with classical immunohistological methods. In these, the specific recognition is provided by the primary antibody (polyclonal or monoclonal) but the secondary detection system can utilize fluorescent, enzyme, or other conjugated secondary antibodies. As a result, an immunohistological staining of tissue section for pathological examination is obtained. Tissues can also be extracted, *e.g.*, with urea and neutral detergent, for the liberation of *Staphylococcus* polypeptides for Western-blot or dot/slot assay. *See, e.g.*, Jalkanen, M. et al. (1985) J. Cell. Biol. 101:976-985; Jalkanen, M. et al. (1987) J. Cell. Biol. 105:3087-3096. In this technique, which is based on the use of cationic solid phases, quantitation of a *Staphylococcus* polypeptide can be accomplished using an isolated *Staphylococcus* polypeptide as a standard. This technique can also be applied to body fluids.

[0315] Other antibody-based methods useful for detecting *Staphylococcus* polypeptide gene expression include immunoassays, such as the ELISA and the radioimmunoassay (RIA). For example, a *Staphylococcus* polypeptide-specific monoclonal antibodies can be used both as an immunoabsorbent and as an enzyme-labeled probe to detect and quantify a *Staphylococcus* polypeptide. The amount of a *Staphylococcus* polypeptide present in the sample can be calculated by reference to the amount present in a standard preparation using a linear regression computer algorithm. Such an ELISA is described in Iacobelli et al. (1988) Breast

Cancer Research and Treatment 11:19-30. In another ELISA assay, two distinct specific monoclonal antibodies can be used to detect *Staphylococcus* polypeptides in a body fluid. In this assay, one of the antibodies is used as the immunoabsorbent and the other as the enzyme-labeled probe.

[0316] The above techniques may be conducted essentially as a "one-step" or "two-step" assay. The "one-step" assay involves contacting the *Staphylococcus* polypeptide with immobilized antibody and, without washing, contacting the mixture with the labeled antibody. The "two-step" assay involves washing before contacting the mixture with the labeled antibody. Other conventional methods may also be employed as suitable. It is usually desirable to immobilize one component of the assay system on a support, thereby allowing other components of the system to be brought into contact with the component and readily removed from the sample. Variations of the above and other immunological methods included in the present invention can also be found in Harlow et al., ANTIBODIES: A LABORATORY MANUAL, (Cold Spring Harbor Laboratory Press, 2nd ed. 1988).

[0317] Suitable enzyme labels include, for example, those from the oxidase group, which catalyze the production of hydrogen peroxide by reacting with substrate. Glucose oxidase is particularly preferred as it has good stability and its substrate (glucose) is readily available. Activity of an oxidase label may be assayed by measuring the concentration of hydrogen peroxide formed by the enzyme-labeled antibody/substrate reaction. Besides enzymes, other suitable labels include radioisotopes, such as iodine ( $^{125}\text{I}$ ,  $^{121}\text{I}$ ), carbon ( $^{14}\text{C}$ ), sulphur ( $^{35}\text{S}$ ), tritium ( $^3\text{H}$ ), indium ( $^{112}\text{In}$ ), and technetium ( $^{99\text{m}}\text{Tc}$ ), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

[0318] Further suitable labels for the *Staphylococcus* polypeptide-specific antibodies of the present invention are provided below. Examples of suitable enzyme labels include malate dehydrogenase, *Staphylococcus* nuclease, delta-5-steroid isomerase, yeast-alcohol dehydrogenase, alpha-glycerol phosphate dehydrogenase, triose phosphate isomerase, peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, beta-galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase, glucoamylase, and acetylcholine esterase.

[0319] Examples of suitable radioisotopic labels include  $^3\text{H}$ ,  $^{111}\text{In}$ ,  $^{125}\text{I}$ ,  $^{131}\text{I}$ ,  $^{32}\text{P}$ ,  $^{35}\text{S}$ ,  $^{14}\text{C}$ ,  $^{51}\text{Cr}$ ,  $^{57}\text{To}$ ,  $^{58}\text{Co}$ ,  $^{59}\text{Fe}$ ,  $^{75}\text{Se}$ ,  $^{152}\text{Eu}$ ,  $^{90}\text{Y}$ ,  $^{67}\text{Cu}$ ,  $^{217}\text{Ci}$ ,  $^{211}\text{At}$ ,  $^{212}\text{Pb}$ ,  $^{47}\text{Sc}$ ,  $^{109}\text{Pd}$ , etc.  $^{111}\text{In}$  is a preferred isotope where *in vivo* imaging is used since it avoids the problem of dehalogenation of the  $^{125}\text{I}$  or  $^{131}\text{I}$ -labeled monoclonal antibody by the liver. In addition, this radionucleotide

has a more favorable gamma emission energy for imaging. See, e.g., Perkins et al. (1985) Eur. J. Nucl. Med. 10:296-301; Carasquillo et al. (1987) J. Nucl. Med. 28:281-287. For example,  $^{111}\text{In}$  coupled to monoclonal antibodies with 1-(P-isothiocyanatobenzyl)-DPTA has shown little uptake in non-tumors tissues, particularly the liver, and therefore enhances specificity of tumor localization. See, Esteban et al. (1987) J. Nucl. Med. 28:861-870.

[0320] Examples of suitable non-radioactive isotopic labels include  $^{157}\text{Gd}$ ,  $^{55}\text{Mn}$ ,  $^{162}\text{Dy}$ ,  $^{52}\text{Tr}$ , and  $^{56}\text{Fe}$ .

[0321] Examples of suitable fluorescent labels include an  $^{152}\text{Eu}$  label, a fluorescein label, an isothiocyanate label, a rhodamine label, a phycoerythrin label, a phycocyanin label, an allophycocyanin label, an o-phthaldehyde label, and a fluorescamine label.

[0322] Examples of suitable toxin labels include, *Pseudomonas* toxin, diphtheria toxin, ricin, and cholera toxin.

[0323] Examples of chemiluminescent labels include a luminal label, an isoluminal label, an aromatic acridinium ester label, an imidazole label, an acridinium salt label, an oxalate ester label, a luciferin label, a luciferase label, and an aequorin label.

[0324] Examples of nuclear magnetic resonance contrasting agents include heavy metal nuclei such as Gd, Mn, and iron.

[0325] Typical techniques for binding the above-described labels to antibodies are provided by Kennedy et al. (1976) Clin. Chim. Acta 70:1-31, and Schurs et al. (1977) Clin. Chim. Acta 81:1-40. Coupling techniques mentioned in the latter are the glutaraldehyde method, the periodate method, the dimaleimide method, the m-maleimidobenzyl-N-hydroxy-succinimide ester method, all of which methods are incorporated by reference herein.

[0326] In a related aspect, the invention includes a diagnostic kit for use in screening serum containing antibodies specific against *S. aureus* infection. Such a kit may include an isolated *S. aureus* antigen comprising an epitope which is specifically immunoreactive with at least one anti-*S. aureus* antibody. Such a kit also includes means for detecting the binding of said antibody to the antigen. In specific embodiments, the kit may include a recombinantly produced or chemically synthesized peptide or polypeptide antigen. The peptide or polypeptide antigen may be attached to a solid support.

[0327] In a more specific embodiment, the detecting means of the above-described kit includes a solid support to which said peptide or polypeptide antigen is attached. Such a kit may also include a non-attached reporter-labeled anti-human antibody. In this embodiment,

binding of the antibody to the *S. aureus* antigen can be detected by binding of the reporter labeled antibody to the anti-*S. aureus* polypeptide antibody.

[0328] In a related aspect, the invention includes a method of detecting *S. aureus* infection in a subject. This detection method includes reacting a body fluid, preferably serum, from the subject with an isolated *S. aureus* antigen, and examining the antigen for the presence of bound antibody. In a specific embodiment, the method includes a polypeptide antigen attached to a solid support, and serum is reacted with the support. Subsequently, the support is reacted with a reporter-labeled anti-human antibody. The support is then examined for the presence of reporter-labeled antibody.

[0329] The solid surface reagent employed in the above assays and kits is prepared by known techniques for attaching protein material to solid support material, such as polymeric beads, dip sticks, 96-well plates or filter material. These attachment methods generally include non-specific adsorption of the protein to the support or covalent attachment of the protein, typically through a free amine group, to a chemically reactive group on the solid support, such as an activated carboxyl, hydroxyl, or aldehyde group. Alternatively, streptavidin coated plates can be used in conjunction with biotinylated antigen(s).

[0330] The polypeptides and antibodies of the present invention, including fragments thereof, may be used to detect Staphylococcus species including *S. aureus* using bio chip and biosensor technology. Bio chip and biosensors of the present invention may comprise the polypeptides of the present invention to detect antibodies, which specifically recognize Staphylococcus species, including *S. aureus*. Bio chip and biosensors of the present invention may also comprise antibodies which specifically recognize the polypeptides of the present invention to detect Staphylococcus species, including *S. aureus* or specific polypeptides of the present invention. Bio chips or biosensors comprising polypeptides or antibodies of the present invention may be used to detect Staphylococcus species, including *S. aureus*, in biological and environmental samples and to diagnose an animal, including humans, with an *S. aureus* or other Staphylococcus infection. Thus, the present invention includes both bio chips and biosensors comprising polypeptides or antibodies of the present invention and methods of their use.

[0331] The bio chips of the present invention may further comprise polypeptide sequences of other pathogens including bacteria, viral, parasitic, and fungal polypeptide sequences, in addition to the polypeptide sequences of the present invention, for use in rapid differential pathogenic detection and diagnosis. The bio chips of the present invention may further

comprise antibodies or fragments thereof specific for other pathogens including bacteria, viral, parasitic, and fungal polypeptide sequences, in addition to the antibodies or fragments thereof of the present invention, for use in rapid differential pathogenic detection and diagnosis. The bio chips and biosensors of the present invention may also be used to monitor an *S. aureus* or other Staphylococcus infection and to monitor the genetic changes (amino acid deletions, insertions, substitutions, etc.) in response to drug therapy in the clinic and drug development in the laboratory. The bio chip and biosensors comprising polypeptides or antibodies of the present invention may also be used to simultaneously monitor the expression of a multiplicity of polypeptides, including those of the present invention. The polypeptides used to comprise a bio chip or biosensor of the present invention may be specified in the same manner as for the fragments, i.e., by their N-terminal and C-terminal positions or length in contiguous amino acid residue. Methods and particular uses of the polypeptides and antibodies of the present invention to detect Staphylococcus species, including *S. aureus*, or specific polypeptides using bio chip and biosensor technology include those known in the art, those of the U.S. Patent Nos. and World Patent Nos. listed above for bio chips and biosensors using polynucleotides of the present invention, and those of: U.S. Patent Nos. 5658732, 5135852, 5567301, 5677196, 5690894 and World Patent Nos. WO9729366, WO9612957, each incorporated herein in their entireties.

**[0332]     *Treatment***

**[0333]     *Agonists and Antagonists - Assays and Molecules***

**[0334]**     The invention also provides a method of screening compounds to identify those which enhance or block the biological activity of the *S. aureus* polypeptides of the present invention. The present invention further provides where the compounds kill or slow the growth of *S. aureus*. The ability of *S. aureus* antagonists, including *S. aureus* ligands, to prophylactically or therapeutically block antibiotic resistance may be easily tested by the skilled artisan. *See, e.g.,* Straden et al. (1997) J Bacteriol. 179(1):9-16.

**[0335]**     An agonist is a compound which increases the natural biological function or which functions in a manner similar to the polypeptides of the present invention, while antagonists decrease or eliminate such functions. Potential antagonists include small organic molecules, peptides, polypeptides, and antibodies that bind to a polypeptide of the invention and thereby inhibit or extinguish its activity.

**[0336]**     The antagonists may be employed for instance to inhibit peptidoglycan cross bridge



formation. Antibodies against *S. aureus* may be employed to bind to and inhibit *S. aureus* activity to treat antibiotic resistance. Any of the above antagonists may be employed in a composition with a pharmaceutically acceptable carrier.

**[0337] Vaccines**

**[0338]** The present invention also provides vaccines comprising one or more polypeptides of the present invention. Heterogeneity in the composition of a vaccine may be provided by combining *S. aureus* polypeptides of the present invention. Multi-component vaccines of this type are desirable because they are likely to be more effective in eliciting protective immune responses against multiple species and strains of the *Staphylococcus* genus than single polypeptide vaccines.

**[0339]** Multi-component vaccines are known in the art to elicit antibody production to numerous immunogenic components. *See, e.g.*, Decker et al. (1996) J. Infect. Dis. 174:S270-275. In addition, a hepatitis B, diphtheria, tetanus, pertussis tetravalent vaccine has recently been demonstrated to elicit protective levels of antibodies in human infants against all four pathogenic agents. *See, e.g.*, Aristegui, J. et al. (1997) Vaccine 15:7-9.

**[0340]** The present invention in addition to single-component vaccines includes multi-component vaccines. These vaccines comprise more than one polypeptide, immunogen or antigen. Thus, a multi-component vaccine would be a vaccine comprising more than one of the *S. aureus* polypeptides of the present invention.

**[0341]** Further within the scope of the invention are whole cell and whole viral vaccines. Such vaccines may be produced recombinantly and involve the expression of one or more of the *S. aureus* polypeptides described in Table 1. For example, the *S. aureus* polypeptides of the present invention may be either secreted or localized intracellular, on the cell surface, or in the periplasmic space. Further, when a recombinant virus is used, the *S. aureus* polypeptides of the present invention may, for example, be localized in the viral envelope, on the surface of the capsid, or internally within the capsid. Whole cells vaccines which employ cells expressing heterologous proteins are known in the art. *See, e.g.*, Robinson, K. et al. (1997) Nature Biotech. 15:653-657; Sirard, J. et al. (1997) Infect. Immun. 65:2029-2033; Chabalgoity, J. et al. (1997) Infect. Immun. 65:2402-2412. These cells may be administered live or may be killed prior to administration. Chabalgoity, J. et al., *supra*, for example, report the successful use in mice of a live attenuated *Salmonella* vaccine strain which expresses a portion of a platyhelminth fatty acid-binding protein as a fusion protein on its cells surface.

**[0342]** A multi-component vaccine can also be prepared using techniques known in the art by combining one or more *S. aureus* polypeptides of the present invention, or fragments thereof, with additional non-staphylococcal components (*e.g.*, diphtheria toxin or tetanus toxin, and/or other compounds known to elicit an immune response). Such vaccines are useful for eliciting protective immune responses to both members of the *Staphylococcus* genus and non-staphylococcal pathogenic agents.

**[0343]** The vaccines of the present invention also include DNA vaccines. DNA vaccines are currently being developed for a number of infectious diseases. *See, et al.*, Boyer, et al. (1997) Nat. Med. 3:526-532; reviewed in Spier, R. (1996) Vaccine 14:1285-1288. Such DNA vaccines contain a nucleotide sequence encoding one or more *S. aureus* polypeptides of the present invention oriented in a manner that allows for expression of the subject polypeptide. For example, the direct administration of plasmid DNA encoding *B. burgdorferi* OspA has been shown to elicit protective immunity in mice against borrelial challenge. *See*, Luke et al. (1997) J. Infect. Dis. 175:91-97.

**[0344]** The present invention also relates to the administration of a vaccine which is co-administered with a molecule capable of modulating immune responses. Kim et al. (1997) Nature Biotech. 15:641-646, for example, report the enhancement of immune responses produced by DNA immunizations when DNA sequences encoding molecules which stimulate the immune response are co-administered. In a similar fashion, the vaccines of the present invention may be co-administered with either nucleic acids encoding immune modulators or the immune modulators themselves. These immune modulators include granulocyte macrophage colony stimulating factor (GM-CSF) and CD86.

**[0345]** The vaccines of the present invention may be used to confer resistance to staphylococcal infection by either passive or active immunization. When the vaccines of the present invention are used to confer resistance to staphylococcal infection through active immunization, a vaccine of the present invention is administered to an animal to elicit a protective immune response which either prevents or attenuates a staphylococcal infection. When the vaccines of the present invention are used to confer resistance to staphylococcal infection through passive immunization, the vaccine is provided to a host animal (*e.g.*, human, dog, or mouse), and the antisera elicited by this antisera is recovered and directly provided to a recipient suspected of having an infection caused by a member of the *Staphylococcus* genus.

**[0346]** The ability to label antibodies, or fragments of antibodies, with toxin molecules provides an additional method for treating staphylococcal infections when passive

immunization is conducted. In this embodiment, antibodies, or fragments of antibodies, capable of recognizing the *S. aureus* polypeptides disclosed herein, or fragments thereof, as well as other *Staphylococcus* proteins, are labeled with toxin molecules prior to their administration to the patient. When such toxin derivatized antibodies bind to *Staphylococcus* cells, toxin moieties will be localized to these cells and will cause their death.

[0347] The present invention thus concerns and provides a means for preventing or attenuating a staphylococcal infection resulting from organisms which have antigens that are recognized and bound by antisera produced in response to the polypeptides of the present invention. As used herein, a vaccine is said to prevent or attenuate a disease if its administration to an animal results either in the total or partial attenuation (*i.e.*, suppression) of a symptom or condition of the disease, or in the total or partial immunity of the animal to the disease.

[0348] The administration of the vaccine (or the antisera which it elicits) may be for either a "prophylactic" or "therapeutic" purpose. When provided prophylactically, the compound(s) are provided in advance of any symptoms of staphylococcal infection. The prophylactic administration of the compound(s) serves to prevent or attenuate any subsequent infection. When provided therapeutically, the compound(s) is provided upon or after the detection of symptoms which indicate that an animal may be infected with a member of the *Staphylococcus* genus. The therapeutic administration of the compound(s) serves to attenuate any actual infection. Thus, the *S. aureus* polypeptides, and fragments thereof, of the present invention may be provided either prior to the onset of infection (so as to prevent or attenuate an anticipated infection) or after the initiation of an actual infection.

[0349] The polypeptides of the invention, whether encoding a portion of a native protein or a functional derivative thereof, may be administered in pure form or may be coupled to a macromolecular carrier. Example of such carriers are proteins and carbohydrates. Suitable proteins which may act as macromolecular carrier for enhancing the immunogenicity of the polypeptides of the present invention include keyhole limpet hemacyanin (KLH) tetanus toxoid, pertussis toxin, bovine serum albumin, and ovalbumin. Methods for coupling the polypeptides of the present invention to such macromolecular carriers are disclosed in Harlow et al., ANTIBODIES: A LABORATORY MANUAL, (Cold Spring Harbor Laboratory Press, 2nd ed. 1988).

[0350] A composition is said to be "pharmacologically or physiologically acceptable" if its administration can be tolerated by a recipient animal and is otherwise suitable for

administration to that animal. Such an agent is said to be administered in a "therapeutically effective amount" if the amount administered is physiologically significant. An agent is physiologically significant if its presence results in a detectable change in the physiology of a recipient patient.

**[0351]** While in all instances the vaccine of the present invention is administered as a pharmacologically acceptable compound, one skilled in the art would recognize that the composition of a pharmacologically acceptable compound varies with the animal to which it is administered. For example, a vaccine intended for human use will generally not be co-administered with Freund's adjuvant. Further, the level of purity of the *S. aureus* polypeptides of the present invention will normally be higher when administered to a human than when administered to a non-human animal.

**[0352]** As would be understood by one of ordinary skill in the art, when the vaccine of the present invention is provided to an animal, it may be in a composition which may contain salts, buffers, adjuvants, or other substances which are desirable for improving the efficacy of the composition. Adjuvants are substances that can be used to specifically augment a specific immune response. These substances generally perform two functions: (1) they protect the antigen(s) from being rapidly catabolized after administration and (2) they nonspecifically stimulate immune responses.

**[0353]** Normally, the adjuvant and the composition are mixed prior to presentation to the immune system, or presented separately, but into the same site of the animal being immunized. Adjuvants can be loosely divided into several groups based upon their composition. These groups include oil adjuvants (for example, Freund's complete and incomplete), mineral salts (for example,  $\text{AlK}(\text{SO}_4)_2$ ,  $\text{AlNa}(\text{SO}_4)_2$ ,  $\text{AlNH}_4(\text{SO}_4)$ , silica, kaolin, and carbon), polynucleotides (for example, poly IC and poly AU acids), and certain natural substances (for example, wax D from *Mycobacterium tuberculosis*, as well as substances found in *Corynebacterium parvum*, or *Bordetella pertussis*, and members of the genus *Brucella*). Other substances useful as adjuvants are the saponins such as, for example, Quil A. (Superfos A/S, Denmark). Preferred adjuvants for use in the present invention include aluminum salts, such as  $\text{AlK}(\text{SO}_4)_2$ ,  $\text{AlNa}(\text{SO}_4)_2$ , and  $\text{AlNH}_4(\text{SO}_4)$ . Examples of materials suitable for use in vaccine compositions are provided in REMINGTON'S PHARMACEUTICAL SCIENCES 1324-1341 (A. Osol, ed, Mack Publishing Co, Easton, PA, (1980) (incorporated herein by reference).

**[0354]** The therapeutic compositions of the present invention can be administered parenterally by injection, rapid infusion, nasopharyngeal absorption (intranasopharyngeally),

dermoabsorption, or orally. The compositions may alternatively be administered intramuscularly, or intravenously. Compositions for parenteral administration include sterile aqueous or non-aqueous solutions, suspensions, and emulsions. Examples of non-aqueous solvents are propylene glycol, polyethylene glycol, vegetable oils such as olive oil, and injectable organic esters such as ethyl oleate. Carriers or occlusive dressings can be used to increase skin permeability and enhance antigen absorption. Liquid dosage forms for oral administration may generally comprise a liposome solution containing the liquid dosage form. Suitable forms for suspending liposomes include emulsions, suspensions, solutions, syrups, and elixirs containing inert diluents commonly used in the art, such as purified water. Besides the inert diluents, such compositions can also include adjuvants, wetting agents, emulsifying and suspending agents, or sweetening, flavoring, or perfuming agents.

[0355] Therapeutic compositions of the present invention can also be administered in encapsulated form. For example, intranasal immunization using vaccines encapsulated in biodegradable microsphere composed of poly(DL-lactide-co-glycolide). *See*, Shahin, R. et al. (1995) *Infect. Immun.* 63:1195-1200. Similarly, orally administered encapsulated *Salmonella typhimurium* antigens can also be used. Allaoui-Attarki, K. et al. (1997) *Infect. Immun.* 65:853-857. Encapsulated vaccines of the present invention can be administered by a variety of routes including those involving contacting the vaccine with mucous membranes (*e.g.*, intranasally, intracolonicly, intraduodenally).

[0356] Many different techniques exist for the timing of the immunizations when a multiple administration regimen is utilized. It is possible to use the compositions of the invention more than once to increase the levels and diversities of expression of the immunoglobulin repertoire expressed by the immunized animal. Typically, if multiple immunizations are given, they will be given one to two months apart.

[0357] According to the present invention, an "effective amount" of a therapeutic composition is one which is sufficient to achieve a desired biological effect. Generally, the dosage needed to provide an effective amount of the composition will vary depending upon such factors as the animal's or human's age, condition, sex, and extent of disease, if any, and other variables which can be adjusted by one of ordinary skill in the art.

[0358] The antigenic preparations of the invention can be administered by either single or multiple dosages of an effective amount. Effective amounts of the compositions of the invention can vary from 0.01-1,000  $\mu\text{g/ml}$  per dose, more preferably 0.1-500  $\mu\text{g/ml}$  per dose, and most preferably 10-300  $\mu\text{g/ml}$  per dose.

**[0359] Examples**

**[0360] *Example 1: Isolation of a Selected DNA Clone From the Deposited Sample***

**[0361]** Three approaches can be used to isolate a *S. aureus* clone comprising a polynucleotide of the present invention from any *S. aureus* genomic DNA library. The *S. aureus* strain ISP3 has been deposited as a convenient source for obtaining a *S. aureus* strain although a wide variety of strains *S. aureus* strains can be used which are known in the art.

**[0362]** *S. aureus* genomic DNA is prepared using the following method. A 20ml overnight bacterial culture grown in a rich medium (e.g., Trypticase Soy Broth, Brain Heart Infusion broth or Super broth), pelleted, washed two times with TES (30mM Tris-pH 8.0, 25mM EDTA, 50mM NaCl), and resuspended in 5ml high salt TES (2.5M NaCl). Lysostaphin is added to final concentration of approx 50ug/ml and the mixture is rotated slowly 1 hour at 37C to make protoplast cells. The solution is then placed in incubator (or place in a shaking water bath) and warmed to 55C. Five hundred micro liter of 20% sarcosyl in TES (final concentration 2%) is then added to lyse the cells. Next, guanidine HCl is added to a final concentration of 7M (3.69g in 5.5 ml). The mixture is swirled slowly at 55C for 60-90 min (solution should clear). A CsCl gradient is then set up in SW41 ultra clear tubes using 2.0ml 5.7M CsCl and overlaying with 2.85M CsCl. The gradient is carefully overlayed with the DNA-containing GuHCl solution. The gradient is spun at 30,000 rpm, 20C for 24 hr and the lower DNA band is collected. The volume is increased to 5 ml with TE buffer. The DNA is then treated with protease K (10 ug/ml) overnight at 37 C, and precipitated with ethanol. The precipitated DNA is resuspended in a desired buffer.

**[0363]** In the first method, a plasmid is directly isolated by screening a plasmid *S. aureus* genomic DNA library using a polynucleotide probe corresponding to a polynucleotide of the present invention. Particularly, a specific polynucleotide with 30-40 nucleotides is synthesized using an Applied Biosystems DNA synthesizer according to the sequence reported. The oligonucleotide is labeled, for instance, with <sup>32</sup>P-γ-ATP using T4 polynucleotide kinase and purified according to routine methods. (See, e.g., Maniatis et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring, NY (1982).) The library is transformed into a suitable host, as indicated above (such as XL-1 Blue (Stratagene)) using techniques known to those of skill in the art. See, e.g., Sambrook et al. *MOLECULAR CLONING: A LABORATORY MANUAL* (Cold Spring Harbor, N.Y. 2nd ed. 1989); Ausubel et al., *CURRENT PROTOCOLS IN MOLECULAR BIOLOGY* (John Wiley and Sons, N.Y.

1989). The transformants are plated on 1.5% agar plates (containing the appropriate selection agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate. These plates are screened using Nylon membranes according to routine methods for bacterial colony screening. See, e.g., Sambrook et al. MOLECULAR CLONING: A LABORATORY MANUAL (Cold Spring Harbor, N.Y. 2nd ed. 1989); Ausubel et al., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY (John Wiley and Sons, N.Y. 1989) or other techniques known to those of skill in the art.

[0364] Alternatively, two primers of 15-25 nucleotides derived from the 5' and 3' ends of a polynucleotide of Table 1 are synthesized and used to amplify the desired DNA by PCR using a *S. aureus* genomic DNA prep (e.g., the deposited *S. aureus* ISP3) as a template. PCR is carried out under routine conditions, for instance, in 25 µl of reaction mixture with 0.5 ug of the above DNA template. A convenient reaction mixture is 1.5-5 mM MgCl<sub>2</sub>, 0.01% (w/v) gelatin, 20 µM each of dATP, dCTP, dGTP, dTTP, 25 pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR (denaturation at 94°C for 1 min; annealing at 55°C for 1 min; elongation at 72°C for 1 min) are performed with a Perkin-Elmer Cetus automated thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and the DNA band with expected molecular weight is excised and purified. The PCR product is verified to be the selected sequence by subcloning and sequencing the DNA product.

[0365] Finally, overlapping oligos of the DNA sequences of Table 1 can be synthesized and used to generate a nucleotide sequence of desired length using PCR methods known in the art.

[0366] ***Example 2(a): Expression and Purification staphylococcal polypeptides in E. coli***

[0367] The bacterial expression vector pQE60 is used for bacterial expression in this example. (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311). pQE60 encodes ampicillin antibiotic resistance ("Ampr") and contains a bacterial origin of replication ("ori"), an IPTG inducible promoter, a ribosome binding site ("RBS"), six codons encoding histidine residues that allow affinity purification using nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin (QIAGEN, Inc., *supra*) and suitable single restriction enzyme cleavage sites. These elements are arranged such that an inserted DNA fragment encoding a polypeptide expresses that polypeptide with the six His residues (i.e., a "6 X His tag") covalently linked to the carboxyl terminus of that polypeptide.

[0368] The DNA sequence encoding the desired portion of a *S. aureus* protein of the

present invention is amplified from *S. aureus* genomic DNA or from the deposited DNA clone using PCR oligonucleotide primers which anneal to the 5' and 3' sequences coding for the portion of the *S. aureus* polynucleotide. Additional nucleotides containing restriction sites to facilitate cloning in the pQE60 vector are added to the 5' and 3' sequences, respectively.

[0369] For cloning the mature protein, the 5' primer has a sequence containing an appropriate restriction site followed by nucleotides of the amino terminal coding sequence of the desired *S. aureus* polynucleotide sequence in Table 1. One of ordinary skill in the art would appreciate that the point in the protein coding sequence where the 5' and 3' primers begin may be varied to amplify a DNA segment encoding any desired portion of the complete protein shorter or longer than the mature form. The 3' primer has a sequence containing an appropriate restriction site followed by nucleotides complementary to the 3' end of the desired coding sequence of Table 1, excluding a stop codon, with the coding sequence aligned with the restriction site so as to maintain its reading frame with that of the six His codons in the pQE60 vector.

[0370] The amplified *S. aureus* DNA fragment and the vector pQE60 are digested with restriction enzymes which recognize the sites in the primers and the digested DNAs are then ligated together. The *S. aureus* DNA is inserted into the restricted pQE60 vector in a manner which places the *S. aureus* protein coding region downstream from the IPTG-inducible promoter and in-frame with an initiating AUG and the six histidine codons.

[0371] The ligation mixture is transformed into competent *E. coli* cells using standard procedures such as those described by Sambrook et al., *supra*. *E. coli* strain M15/rep4, containing multiple copies of the plasmid pREP4, which expresses the lac repressor and confers kanamycin resistance ("Kanr"), is used in carrying out the illustrative example described herein. This strain, which is only one of many that are suitable for expressing a *S. aureus* polypeptide, is available commercially (QIAGEN, Inc., *supra*). Transformants are identified by their ability to grow on LB plates in the presence of ampicillin and kanamycin. Plasmid DNA is isolated from resistant colonies and the identity of the cloned DNA confirmed by restriction analysis, PCR and DNA sequencing.

[0372] Clones containing the desired constructs are grown overnight ("O/N") in liquid culture in LB media supplemented with both ampicillin (100 µg/ml) and kanamycin (25 µg/ml). The O/N culture is used to inoculate a large culture, at a dilution of approximately 1:25 to 1:250. The cells are grown to an optical density at 600 nm ("OD600") of between 0.4 and 0.6. Isopropyl-β-D-thiogalactopyranoside ("IPTG") is then added to a final concentration



of 1 mM to induce transcription from the lac repressor sensitive promoter, by inactivating the lacI repressor. Cells subsequently are incubated further for 3 to 4 hours. Cells then are harvested by centrifugation.

**[0373]** The cells are then stirred for 3-4 hours at 4°C in 6M guanidine-HCl, pH 8. The cell debris is removed by centrifugation, and the supernatant containing the *S. aureus* polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (QIAGEN, Inc., *supra*). Proteins with a 6 x His tag bind to the Ni-NTA resin with high affinity and can be purified in a simple one-step procedure (for details see: The QIAexpressionist, 1995, QIAGEN, Inc., *supra*). Briefly the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the *S. aureus* polypeptide is eluted with 6 M guanidine-HCl, pH 5.

**[0374]** The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be successfully refolded while immobilized on the Ni-NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins can be eluted by the addition of 250 mM imidazole. Imidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4° C or frozen at -80° C.

**[0375]** Alternatively, the polypeptides of the present invention can be produced by a non-denaturing method. In this method, after the cells are harvested by centrifugation, the cell pellet from each liter of culture is resuspended in 25 ml of Lysis Buffer A at 4°C (Lysis Buffer A = 50 mM Na-phosphate, 300 mM NaCl, 10 mM 2-mercaptoethanol, 10% Glycerol, pH 7.5 with 1 tablet of Complete EDTA-free protease inhibitor cocktail (Boehringer Mannheim #1873580) per 50 ml of buffer). Absorbance at 550 nm is approximately 10-20 O.D./ml. The suspension is then put through three freeze/thaw cycles from -70°C (using a ethanol-dry ice bath) up to room temperature. The cells are lysed via sonication in short 10 sec bursts over 3 minutes at approximately 80W while kept on ice. The sonicated sample is then centrifuged at 15,000 RPM for 30 minutes at 4°C. The supernatant is passed through a column containing 1.0 ml of CL-4B resin to pre-clear the sample of any proteins that may bind to agarose non-specifically, and the flow-through fraction is collected.

[0376] The pre-cleared flow-through is applied to a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (Quiagen, Inc., *supra*). Proteins with a 6 X His tag bind to the Ni-NTA resin with high affinity and can be purified in a simple one-step procedure. Briefly, the supernatant is loaded onto the column in Lysis Buffer A at 4°C, the column is first washed with 10 volumes of Lysis Buffer A until the A280 of the eluate returns to the baseline. Then, the column is washed with 5 volumes of 40 mM Imidazole (92% Lysis Buffer A / 8% Buffer B) (Buffer B = 50 mM Na-Phosphate, 300 mM NaCl, 10% Glycerol, 10 mM 2-mercaptoethanol, 500 mM Imidazole, pH of the final buffer should be 7.5). The protein is eluted off of the column with a series of increasing Imidazole solutions made by adjusting the ratios of Lysis Buffer A to Buffer B. Three different concentrations are used: 3 volumes of 75 mM Imidazole, 3 volumes of 150 mM Imidazole, 5 volumes of 500 mM Imidazole. The fractions containing the purified protein are analyzed using 8 %, 10 % or 14% SDS-PAGE depending on the protein size. The purified protein is then dialyzed 2X against phosphate-buffered saline (PBS) in order to place it into an easily workable buffer. The purified protein is stored at 4°C or frozen at -80°

[0377] The following is another alternative method may be used to purify *S. aureus* expressed in *E coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

[0378] Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells are harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

[0379] The cells are then lysed by passing the solution through a microfluidizer (Microfluidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 x g for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

[0380] The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 x g centrifugation for 15 min., the pellet is discarded and the *S. aureus* polypeptide-containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

[0381] Following high speed centrifugation (30,000 x g) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

[0382] To clarify the refolded *S. aureus* polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 µm membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

[0383] Fractions containing the *S. aureus* polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A<sub>280</sub> monitoring of the effluent. Fractions containing the *S. aureus* polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

[0384] The resultant *S. aureus* polypeptide exhibits greater than 95% purity after the above refolding and purification steps. No major contaminant bands are observed from Coomassie blue stained 16% SDS-PAGE gel when 5 µg of purified protein is loaded. The purified protein is also tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

[0385] ***Example 2(b): Expression and Purification staphylococcal polypeptides in E. coli***

[0386] Alternatively, the vector pQE10 can be used to clone and express polypeptides of the present invention. The difference being such that an inserted DNA fragment encoding a polypeptide expresses that polypeptide with the six His residues (i.e., a "6 X His tag") covalently linked to the amino terminus of that polypeptide. The bacterial expression vector

pQE10 (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311) is used in this example . The components of the pQE10 plasmid are arranged such that the inserted DNA sequence encoding a polypeptide of the present invention expresses the polypeptide with the six His residues (*i.e.*, a "6 X His tag")) covalently linked to the amino terminus.

[0387] The DNA sequences encoding the desired portions of a polypeptide of Table 1 are amplified using PCR oligonucleotide primers from either genomic *S. aureus* DNA or DNA from the plasmid clones listed in Table 1 clones of the present invention. The PCR primers anneal to the nucleotide sequences encoding the desired amino acid sequence of a polypeptide of the present invention. Additional nucleotides containing restriction sites to facilitate cloning in the pQE10 vector are added to the 5' and 3' primer sequences, respectively.

[0388] For cloning a polypeptide of the present invention, the 5' and 3' primers are selected to amplify their respective nucleotide coding sequences. One of ordinary skill in the art would appreciate that the point in the protein coding sequence where the 5' and 3' primers begins may be varied to amplify a DNA segment encoding any desired portion of a polypeptide of the present invention. The 5' primer is designed so the coding sequence of the 6 X His tag is aligned with the restriction site so as to maintain its reading frame with that of *S. aureus* polypeptide. The 3' is designed to include an stop codon. The amplified DNA fragment is then cloned, and the protein expressed, as described above for the pQE60 plasmid.

[0389] The DNA sequences encoding the amino acid sequences of Table 1 may also be cloned and expressed as fusion proteins by a protocol similar to that described directly above, wherein the pET-32b(+) vector (Novagen, 601 Science Drive, Madison, WI 53711) is preferentially used in place of pQE10.

**[0390] Example 2(c): Expression and Purification of Staphylococcus polypeptides in *E. coli***

[0391] The bacterial expression vector pQE60 is used for bacterial expression in this example (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311). However, in this example, the polypeptide coding sequence is inserted such that translation of the six His codons is prevented and, therefore, the polypeptide is produced with no 6 X His tag.

[0392] The DNA sequence encoding the desired portion of the *S. aureus* amino acid sequence is amplified from a *S. aureus* genomic DNA prep using PCR oligonucleotide primers which anneal to the 5' and 3' nucleotide sequences corresponding to the desired portion of the *S. aureus* polypeptides. Additional nucleotides containing restriction sites to facilitate cloning

in the pQE60 vector are added to the 5' and 3' primer sequences.

**[0393]** For cloning a *S. aureus* polypeptides of the present invention, 5' and 3' primers are selected to amplify their respective nucleotide coding sequences. One of ordinary skill in the art would appreciate that the point in the protein coding sequence where the 5' and 3' primers begin may be varied to amplify a DNA segment encoding any desired portion of a polypeptide of the present invention. The 3' and 5' primers contain appropriate restriction sites followed by nucleotides complementary to the 5' and 3' ends of the coding sequence respectively. The 3' primer is additionally designed to include an in-frame stop codon.

**[0394]** The amplified *S. aureus* DNA fragments and the vector pQE60 are digested with restriction enzymes recognizing the sites in the primers and the digested DNAs are then ligated together. Insertion of the *S. aureus* DNA into the restricted pQE60 vector places the *S. aureus* protein coding region including its associated stop codon downstream from the IPTG-inducible promoter and in-frame with an initiating AUG. The associated stop codon prevents translation of the six histidine codons downstream of the insertion point.

**[0395]** The ligation mixture is transformed into competent *E. coli* cells using standard procedures such as those described by Sambrook et al. *E. coli* strain M15/rep4, containing multiple copies of the plasmid pREP4, which expresses the lac repressor and confers kanamycin resistance ("Kanr"), is used in carrying out the illustrative example described herein. This strain, which is only one of many that are suitable for expressing *S. aureus* polypeptide, is available commercially (QIAGEN, Inc., *supra*). Transformants are identified by their ability to grow on LB plates in the presence of ampicillin and kanamycin. Plasmid DNA is isolated from resistant colonies and the identity of the cloned DNA confirmed by restriction analysis, PCR and DNA sequencing.

**[0396]** Clones containing the desired constructs are grown overnight ("O/N") in liquid culture in LB media supplemented with both ampicillin (100 µg/ml) and kanamycin (25 µg/ml). The O/N culture is used to inoculate a large culture, at a dilution of approximately 1:25 to 1:250. The cells are grown to an optical density at 600 nm ("OD600") of between 0.4 and 0.6. isopropyl-b-D-thiogalactopyranoside ("IPTG") is then added to a final concentration of 1 mM to induce transcription from the *lac* repressor sensitive promoter, by inactivating the lacI repressor. Cells subsequently are incubated further for 3 to 4 hours. Cells then are harvested by centrifugation.

**[0397]** To purify the *S. aureus* polypeptide, the cells are then stirred for 3-4 hours at 4°C in 6M guanidine-HCl, pH 8. The cell debris is removed by centrifugation, and the supernatant

containing the *S. aureus* polypeptide is dialyzed against 50 mM Na-acetate buffer pH 6, supplemented with 200 mM NaCl. Alternatively, the protein can be successfully refolded by dialyzing it against 500 mM NaCl, 20% glycerol, 25 mM Tris/HCl pH 7.4, containing protease inhibitors. After renaturation the protein can be purified by ion exchange, hydrophobic interaction and size exclusion chromatography. Alternatively, an affinity chromatography step such as an antibody column can be used to obtain pure *S. aureus* polypeptide. The purified protein is stored at 4° C or frozen at -80° C.

**[0398]** The following alternative method may be used to purify *S. aureus* polypeptides expressed in *E. coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

**[0399]** Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells are harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

**[0400]** The cells were then lysed by passing the solution through a microfluidizer (Microfluidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 x g for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

**[0401]** The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 x g centrifugation for 15 min., the pellet is discarded and the *S. aureus* polypeptide-containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

**[0402]** Following high speed centrifugation (30,000 x g) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

**[0403]** To clarify the refolded *S. aureus* polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 µm membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered

sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

[0404] Fractions containing the *S. aureus* polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A<sub>280</sub> monitoring of the effluent. Fractions containing the *S. aureus* polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

[0405] The resultant *S. aureus* polypeptide exhibits greater than 95% purity after the above refolding and purification steps. No major contaminant bands are observed from Coomassie blue stained 16% SDS-PAGE gel when 5 µg of purified protein is loaded. The purified protein is also tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

[0406] ***Example 2(d): Cloning and Expression of S. aureus in Other Bacteria***

[0407] *S. aureus* polypeptides also can be produced in: *S. aureus* using the methods of S. Skinner et al., (1988) Mol. Microbiol. 2:289-297 or J. I. Moreno (1996) Protein Expr. Purif. 8(3):332-340; *Lactobacillus* using the methods of C. Rush et al., 1997 Appl. Microbiol. Biotechnol. 47(5):537-542; or in *Bacillus subtilis* using the methods Chang et al., U.S. Patent No. 4,952,508.

[0408] ***Example 3: Cloning and Expression in COS Cells***

[0409] A *S. aureus* expression plasmid is made by cloning a portion of the DNA encoding a *S. aureus* polypeptide into the expression vector pDNAI/Amp or pDNAIII (which can be obtained from Invitrogen, Inc.). The expression vector pDNAI/amp contains: (1) an *E. coli* origin of replication effective for propagation in *E. coli* and other prokaryotic cells; (2) an ampicillin resistance gene for selection of plasmid-containing prokaryotic cells; (3) an SV40

origin of replication for propagation in eukaryotic cells; (4) a CMV promoter, a polylinker, an SV40 intron; (5) several codons encoding a hemagglutinin fragment (i.e., an "HA" tag to facilitate purification) followed by a termination codon and polyadenylation signal arranged so that a DNA can be conveniently placed under expression control of the CMV promoter and operably linked to the SV40 intron and the polyadenylation signal by means of restriction sites in the polylinker. The HA tag corresponds to an epitope derived from the influenza hemagglutinin protein described by Wilson et al. 1984 Cell 37:767. The fusion of the HA tag to the target protein allows easy detection and recovery of the recombinant protein with an antibody that recognizes the HA epitope. pDNAIII contains, in addition, the selectable neomycin marker.

[0410] A DNA fragment encoding a *S. aureus* polypeptide is cloned into the polylinker region of the vector so that recombinant protein expression is directed by the CMV promoter. The plasmid construction strategy is as follows. The DNA from a *S. aureus* genomic DNA prep is amplified using primers that contain convenient restriction sites, much as described above for construction of vectors for expression of *S. aureus* in *E. coli*. The 5' primer contains a Kozak sequence, an AUG start codon, and nucleotides of the 5' coding region of the *S. aureus* polypeptide. The 3' primer, contains nucleotides complementary to the 3' coding sequence of the *S. aureus* DNA, a stop codon, and a convenient restriction site.

[0411] The PCR amplified DNA fragment and the vector, pDNAI/Amp, are digested with appropriate restriction enzymes and then ligated. The ligation mixture is transformed into an appropriate *E. coli* strain such as SURE™ (Stratagene Cloning Systems, La Jolla, CA 92037), and the transformed culture is plated on ampicillin media plates which then are incubated to allow growth of ampicillin resistant colonies. Plasmid DNA is isolated from resistant colonies and examined by restriction analysis or other means for the presence of the fragment encoding the *S. aureus* polypeptide

[0412] For expression of a recombinant *S. aureus* polypeptide, COS cells are transfected with an expression vector, as described above, using DEAE-dextran, as described, for instance, by Sambrook et al. (*supra*). Cells are incubated under conditions for expression of *S. aureus* by the vector.

[0413] Expression of the *S. aureus*-HA fusion protein is detected by radiolabeling and immunoprecipitation, using methods described in, for example Harlow et al., *supra*.. To this end, two days after transfection, the cells are labeled by incubation in media containing <sup>35</sup>S-cysteine for 8 hours. The cells and the media are collected, and the cells are washed and the



lysed with detergent-containing RIPA buffer: 150 mM NaCl, 1% NP-40, 0.1% SDS, 1% NP-40, 0.5% DOC, 50 mM TRIS, pH 7.5, as described by Wilson et al. (*supra*). Proteins are precipitated from the cell lysate and from the culture media using an HA-specific monoclonal antibody. The precipitated proteins then are analyzed by SDS-PAGE and autoradiography. An expression product of the expected size is seen in the cell lysate, which is not seen in negative controls.

**[0414]     *Example 4: Cloning and Expression in CHO Cells***

**[0415]**     The vector pC4 is used for the expression of *S. aureus* polypeptide in this example. Plasmid pC4 is a derivative of the plasmid pSV2-dhfr (ATCC Accession No. 37146). The plasmid contains the mouse DHFR gene under control of the SV40 early promoter. Chinese hamster ovary cells or other cells lacking dihydrofolate activity that are transfected with these plasmids can be selected by growing the cells in a selective medium (alpha minus MEM, Life Technologies) supplemented with the chemotherapeutic agent methotrexate. The amplification of the DHFR genes in cells resistant to methotrexate (MTX) has been well documented. *See, e.g.,* Alt et al., 1978, J. Biol. Chem. 253:1357-1370; Hamlin et al., 1990, Biochem. et Biophys. Acta, 1097:107-143; Page et al., 1991, Biotechnology 9:64-68. Cells grown in increasing concentrations of MTX develop resistance to the drug by overproducing the target enzyme, DHFR, as a result of amplification of the DHFR gene. If a second gene is linked to the DHFR gene, it is usually co-amplified and over-expressed. It is known in the art that this approach may be used to develop cell lines carrying more than 1,000 copies of the amplified gene(s). Subsequently, when the methotrexate is withdrawn, cell lines are obtained which contain the amplified gene integrated into one or more chromosome(s) of the host cell.

**[0416]**     Plasmid pC4 contains the strong promoter of the long terminal repeat (LTR) of the Rouse Sarcoma Virus, for expressing a polypeptide of interest, Cullen, et al. (1985) Mol. Cell. Biol. 5:438-447; plus a fragment isolated from the enhancer of the immediate early gene of human cytomegalovirus (CMV), Boshart, et al., 1985, Cell 41:521-530. Downstream of the promoter are the following single restriction enzyme cleavage sites that allow the integration of the genes: *Bam* HI, *Xba* I, and *Asp* 718. Behind these cloning sites the plasmid contains the 3' intron and polyadenylation site of the rat preproinsulin gene. Other high efficiency promoters can also be used for the expression, e.g., the human  $\beta$ -actin promoter, the SV40 early or late promoters or the long terminal repeats from other retroviruses, e.g., HIV and HTLV. Clontech's Tet-Off and Tet-On gene expression systems and similar systems can be used to

express the *S. aureus* polypeptide in a regulated way in mammalian cells (Gossen et al., 1992, Proc. Natl. Acad. Sci. USA 89:5547-5551. For the polyadenylation of the mRNA other signals, e.g., from the human growth hormone or globin genes can be used as well. Stable cell lines carrying a gene of interest integrated into the chromosomes can also be selected upon co-transfection with a selectable marker such as gpt, G418 or hygromycin. It is advantageous to use more than one selectable marker in the beginning, e.g., G418 plus methotrexate.

[0417] The plasmid pC4 is digested with the restriction enzymes and then dephosphorylated using calf intestinal phosphates by procedures known in the art. The vector is then isolated from a 1% agarose gel. The DNA sequence encoding the *S. aureus* polypeptide is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' sequences of the desired portion of the gene. A 5' primer containing a restriction site, a Kozak sequence, an AUG start codon, and nucleotides of the 5' coding region of the *S. aureus* polypeptide is synthesized and used. A 3' primer, containing a restriction site, stop codon, and nucleotides complementary to the 3' coding sequence of the *S. aureus* polypeptides is synthesized and used. The amplified fragment is digested with the restriction endonucleases and then purified again on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC4 using, for instance, restriction enzyme analysis.

[0418] Chinese hamster ovary cells lacking an active DHFR gene are used for transfection. Five µg of the expression plasmid pC4 is cotransfected with 0.5 µg of the plasmid pSVneo using a lipid-mediated transfection agent such as Lipofectin™ or LipofectAMINE™ (LifeTechnologies Gaithersburg, MD). The plasmid pSV2-neo contains a dominant selectable marker, the *neo* gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of methotrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 µM, 2 µM, 5 µM, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100-200 µM. Expression

of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

**[0419]     *Example 5: Quantitative Murine Soft Tissue Infection Model for S. aureus***

**[0420]**     Compositions of the present invention, including polypeptides and peptides, are assayed for their ability to function as vaccines or to enhance/stimulate an immune response to a bacterial species (e.g., *S. aureus*) using the following quantitative murine soft tissue infection model. Mice (e.g., NIH Swiss female mice, approximately 7 weeks old) are first treated with a biologically protective effective amount, or immune enhancing/stimulating effective amount of a composition of the present invention using methods known in the art, such as those discussed above. See, e.g., Harlow et al., ANTIBODIES: A LABORATORY MANUAL, (Cold Spring Harbor Laboratory Press, 2nd ed. 1988). An example of an appropriate starting dose is 20ug per animal.

**[0421]**     The desired bacterial species used to challenge the mice, such as *S. aureus*, is grown as an overnight culture. The culture is diluted to a concentration of  $5 \times 10^8$  cfu/ml, in an appropriate media, mixed well, serially diluted, and titered. The desired doses are further diluted 1:2 with sterilized Cytodex 3 microcarrier beads preswollen in sterile PBS (3g/100ml). Mice are anesthetize briefly until docile, but still mobile and injected with 0.2 ml of the Cytodex 3 bead/bacterial mixture into each animal subcutaneously in the inguinal region. After four days, counting the day of injection as day one, mice are sacrificed and the contents of the abscess is excised and placed in a 15 ml conical tube containing 1.0ml of sterile PBS. The contents of the abscess is then enzymatically treated and plated as follows.

**[0422]**     The abscess is first disrupted by vortexing with sterilized glass beads placed in the tubes. 3.0mls of prepared enzyme mixture (1.0ml Collagenase D (4.0 mg/ml), 1.0ml Trypsin (6.0 mg/ml) and 8.0 ml PBS) is then added to each tube followed by a 20 min. incubation at 37C. The solution is then centrifuged and the supernatant drawn off. 0.5 ml dH2O is then added and the tubes are vortexed and then incubated for 10 min. at room temperature. 0.5 ml media is then added and samples are serially diluted and plated onto agar plates, and grown overnight at 37C. Plates with distinct and separate colonies are then counted, compared to positive and negative control samples, and quantified. The method can be used to identify composition and determine appropriate and effective doses for humans and other animals by comparing the effective doses of compositions of the present invention with compositions known in the art to be effective in both mice and humans. Doses for the effective treatment of

humans and other animals, using compositions of the present invention, are extrapolated using the data from the above experiments of mice. It is appreciated that further studies in humans and other animals may be needed to determine the most effective doses using methods of clinical practice known in the art.

**[0423]     *Example 6: Murine Systemic Neutropenic Model for S. aureus Infection***

**[0424]**     Compositions of the present invention, including polypeptides and peptides, are assayed for their ability to function as vaccines or to enhance/stimulate an immune response to a bacterial species (e.g., *S. aureus*) using the following qualitative murine systemic neutropenic model. In addition, antibodies of the present invention are employed to provide passive immune or immunophylatic therapy prior to or post *S aureus* infection. Mice (e.g., NIH Swiss female mice, approximately 7 weeks old) are first treated with a biologically protective effective amount, or immune enhancing/stimulating effective amount of a composition of the present invention using methods known in the art, such as those discussed above. *See, e.g.*, Harlow et al., ANTIBODIES: A LABORATORY MANUAL, (Cold Spring Harbor Laboratory Press, 2nd ed. 1988). An example of an appropriate starting dose is 20ug of protein per animal. Mice are then injected with 250 - 300 mg/kg cyclophosphamide intraperitoneally. Counting the day of C.P. injection as day one, the mice are left untreated for 5 days to begin recovery of PMNL'S.

**[0425]**     The desired bacterial species used to challenge the mice, such as *S. aureus*, is grown as an overnight culture. The culture is diluted to a concentration of  $5 \times 10^8$  cfu/ml, in an appropriate media, mixed well, serially diluted, and titered. The desired doses are further diluted 1:2 in 4% Brewer's yeast in media.

**[0426]**     Mice are injected with the bacteria/brewer's yeast challenge intraperitoneally. The Brewer's yeast solution alone is used as a control. The mice are then monitored twice daily for the first week following challenge, and once a day for the next week to ascertain morbidity and mortality. Mice remaining at the end of the experiment are sacrificed. The method can be used to identify compositions and determine appropriate and effective doses for humans and other animals by comparing the effective doses of compositions of the present invention with compositions known in the art to be effective in both mice and humans. Doses for the effective treatment of humans and other animals, using compositions of the present invention, are extrapolated using the data from the above experiments of mice. It is appreciated that further

studies in humans and other animals may be needed to determine the most effective doses using methods of clinical practice known in the art.

**[0427]     *Example 7: Murine Lethal Sepsis Model***

**[0428]**     *S. aureus* polypeptides of the present invention can be evaluated for potential vaccine efficacy using the murine lethal sepsis model. In this model, mice are challenged with low lethal doses (for example, between  $10^6$  and  $10^7$  colony forming units [cfu]) of virulent strains of *S. aureus*. Initial studies are conducted to determine a less virulent yet lethal strain of *S. aureus* to determine its LD<sub>50</sub>. Polypeptides of the present invention (e.g., the polypeptides described in Table 1, fragments thereof and fragments that comprise the epitopes shown in Table 4), produced as Examples 2a-d, 3 and 4, and optionally conjugated with another immunogen, are tested as vaccine candidates. Vaccine candidates are selected as protective antigens if they can protect against death when approximately 100 times the LD<sub>50</sub> of the strain is employed. Immunized mice are then challenged with a lethal dose of *S. aureus*.

**[0429]**     More specifically, female C2H/HeJ mice are immunized subcutaneously in groups of 10 with 15 ug of the protein of the present invention formulated in complete Freund's adjuvant (CFA). Twenty one days later, mice are boosted in the same way with protein formulated in incomplete Freund's adjuvant. Twenty-eight days following the boost, animals are bled and immune titers against *S. aureus* proteins are determined by ELISA. 35 days following the boost, a freshly prepared culture of *S. aureus* in BHI (Brain Heart Infusion) both is diluted to approximately 35 to 100xLD<sub>50</sub> in sterile PBS. A lethal dose is then injected intraperitoneally into mice in a volume of 100 ul. Mice are monitored for 14 days for mortality. Survival rate is compared with a sham group immunized with PBS and adjuvant alone.

**[0430]     *Example 8: Identifying Vaccine Antigens Prevalent in S. Aureus Strains***

**[0431]**     It is further determined whether the majority of the most prevalent *S. aureus* strains express the vaccine antigen(s) and polypeptide(s) identified by the lethal model of Example 7 or the models of Examples 5 or 6. Immunoblot analysis is performed with cell lysates prepared from Staphylococcus strains representative of the major capsular serotypes and probed with polyclonal antisera specific for the protective antigens. A preferred vaccine is comprised of a serological epitope of the polypeptide of the present invention that is well conserved among the majority of infective Staphylococcus serotypes.

**[0432]     *Example 9: Production of an Antibody***

***a) Hybridoma Technology***

**[0433]**     The antibodies of the present invention can be prepared by a variety of methods. (See, Current Protocols, Chapter 2.) As one example of such methods, cells expressing polypeptide(s) of the invention are administered to an animal to induce the production of sera containing polyclonal antibodies. In a preferred method, a preparation of polypeptide(s) of the invention is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

**[0434]**     Monoclonal antibodies specific for polypeptide(s) of the invention are prepared using hybridoma technology. (Kohler et al., Nature 256:495 (1975); Kohler et al., Eur. J. Immunol. 6:511 (1976); Kohler et al., Eur. J. Immunol. 6:292 (1976); Hammerling et al., in: Monoclonal Antibodies and T-Cell Hybridomas, Elsevier, N.Y., pp. 563-681 (1981)). In general, an animal (preferably a mouse) is immunized with polypeptide(s) of the invention or, more preferably, with a secreted polypeptide-expressing cell. Such polypeptide-expressing cells are cultured in any suitable tissue culture medium, preferably in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 µg/ml of streptomycin.

**[0435]**     The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP2O), available from the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands et al. (Gastroenterology 80:225-232 (1981)). The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide(s) of the invention.

**[0436]**     Alternatively, additional antibodies capable of binding to polypeptide(s) of the invention can be produced in a two-step procedure using anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a mouse. The

splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the protein-specific antibody can be blocked by polypeptide(s) of the invention. Such antibodies comprise anti-idiotypic antibodies to the protein-specific antibody and are used to immunize an animal to induce formation of further protein-specific antibodies. For in vivo use of antibodies in humans, an antibody is "humanized". Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric and humanized antibodies are known in the art and are discussed herein. (See, for review, Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Cabilly et al., U.S. Patent No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., Nature 312:643 (1984); Neuberger et al., Nature 314:268 (1985).)

***b) Isolation Of Antibody Fragments Directed Against  
Polypeptide(s) From A Library Of scFvs***

[0437] Naturally occurring V-genes isolated from human PBLs are constructed into a library of antibody fragments which contain reactivities against polypeptide(s) of the invention to which the donor may or may not have been exposed (see e.g., U.S. Patent 5,885,793 incorporated herein by reference in its entirety).

**[0438] *Rescue of the Library***

[0439] A library of scFvs is constructed from the RNA of human PBLs as described in PCT publication WO 92/01047. To rescue phage displaying antibody fragments, approximately 10<sup>9</sup> E. coli harboring the phagemid are used to inoculate 50 ml of 2xTY containing 1% glucose and 100 µg/ml of ampicillin (2xTY-AMP-GLU) and grown to an O.D. of 0.8 with shaking. Five ml of this culture is used to inoculate 50 ml of 2xTY-AMP-GLU, 2 x 10<sup>8</sup> TU of delta gene 3 helper (M13 delta gene III, see PCT publication WO 92/01047) are added and the culture incubated at 37°C for 45 minutes without shaking and then at 37°C for 45 minutes with shaking. The culture is centrifuged at 4000 r.p.m. for 10 min. and the pellet resuspended in 2 liters of 2xTY containing 100 µg/ml ampicillin and 50 µg/ml kanamycin and grown overnight. Phage are prepared as described in PCT publication WO 92/01047.

[0440] M13 delta gene III is prepared as follows: M13 delta gene III helper phage does not encode gene III protein, hence the phage(mid) displaying antibody fragments have a greater avidity of binding to antigen. Infectious M13 delta gene III particles are made by growing the

helper phage in cells harboring a pUC19 derivative supplying the wild type gene III protein during phage morphogenesis. The culture is incubated for 1 hour at 37° C without shaking and then for a further hour at 37°C with shaking. Cells are spun down (IEC-Centra 8,400 r.p.m. for 10 min), resuspended in 300 ml 2xTY broth containing 100 µg ampicillin/ml and 25 µg kanamycin/ml (2xTY-AMP-KAN) and grown overnight, shaking at 37°C. Phage particles are purified and concentrated from the culture medium by two PEG-precipitations (Sambrook et al., 1990), resuspended in 2 ml PBS and passed through a 0.45 µm filter (Minisart NML; Sartorius) to give a final concentration of approximately 10<sup>13</sup> transducing units/ml (ampicillin-resistant clones).

**[0441]     *Panning of the Library***

**[0442]**     Immuntubes (Nunc) are coated overnight in PBS with 4 ml of either 100 µg/ml or 10 µg/ml of a polypeptide of the present invention. Tubes are blocked with 2% Marvel-PBS for 2 hours at 37°C and then washed 3 times in PBS. Approximately 10<sup>13</sup> TU of phage is applied to the tube and incubated for 30 minutes at room temperature tumbling on an over and under turntable and then left to stand for another 1.5 hours. Tubes are washed 10 times with PBS 0.1% Tween-20 and 10 times with PBS. Phage are eluted by adding 1 ml of 100 mM triethylamine and rotating 15 minutes on an under and over turntable after which the solution is immediately neutralized with 0.5 ml of 1.0M Tris-HCl, pH 7.4. Phage are then used to infect 10 ml of mid-log E. coli TG1 by incubating eluted phage with bacteria for 30 minutes at 37°C. The E. coli are then plated on TYE plates containing 1% glucose and 100 µg/ml ampicillin. The resulting bacterial library is then rescued with delta gene 3 helper phage as described above to prepare phage for a subsequent round of selection. This process is then repeated for a total of 4 rounds of affinity purification with tube-washing increased to 20 times with PBS, 0.1% Tween-20 and 20 times with PBS for rounds 3 and 4.

**[0443]     *Characterization of Binders***

**[0444]**     Eluted phage from the 3rd and 4th rounds of selection are used to infect E. coli HB 2151 and soluble scFv is produced (Marks, et al., 1991) from single colonies for assay. ELISAs are performed with microtitre plates coated with either 10 µg/ml of the polypeptide of the present invention in 50 mM bicarbonate pH 9.6. Clones positive in ELISA are further characterized by PCR fingerprinting (see, e.g., PCT publication WO 92/01047) and then by sequencing. These ELISA positive clones may also be further characterized by techniques known in the art, such as, for example, epitope mapping, binding affinity, receptor signal transduction, ability to block or competitively inhibit antibody/antigen binding, and



competitive agonistic or antagonistic activity.

**[0445]** The disclosure of all publications (including patents, patent applications, journal articles, laboratory manuals, books, or other documents) cited herein and the sequence listings are hereby incorporated by reference in their entireties.

**[0446]** The present invention is not to be limited in scope by the specific embodiments described herein, which are intended as single illustrations of individual aspects of the invention. Functionally equivalent methods and components are within the scope of the invention, in addition to those shown and described herein and will become apparent to those skilled in the art from the foregoing description and accompanying drawings. Such modifications are intended to fall within the scope of the appended claims.